RADIO ENGINEERING and ELECTRONIC PHYSICS

English Edition of

РАДИОТЕХНИКА И ЭЛЕКТРОНИКА

Published by the American Institute of Electrical Engineers
with the aid of a grant from the National Science Foundation

No.4 April 1961

Translated and Produced by Royer and Roger, Inc.

AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS

Established 1884

345 East Forty-Seventh Street
New York 17, N.Y.

The English edition of RADIO ENGINEERING AND ELECTRONIC PHYSICS is published by the American Institute of Electrical Engineers with the aid of a grant from the National Science Foundation.

Also published under the same arrangement are the Russian electronic journals RADIO ENGINEERING and TELECOMMUNICATIONS.

RADIO ENGINEERING AND ELECTRONIC PHYSICS

(РАДИОТЕХНИКА И ЭЛЕКТРОНИКА)

Publication of the Institute of Radio Engineering and Electronic Physics,
Academy of Sciences of the USSR

Translated and Produced
by
Royer and Roger, Inc.

Subscriptions to Radio Engineering and Electronic Physics should be sent to AIEE
Special Subscription Department
41 East 28th Street, New York 16, New York

1961 Subscription rates:

	$	£
Individuals	28.50	10
Libraries, institutes, govt. agencies	57.00	20

12 issues per annum comprising approximately 1900 pages

ELECTRODYNAMIC THEORY OF A DENSE PLANE GRATING OF PARALLEL CONDUCTORS

A.N. Sivov

In this report is solved the electrodynamic problem of determining the coefficients of reflection and transmission for inclined incidence of a plane electromagnetic wave at a plane grating in free space. The period of the grating is assumed to be small in comparison with the wavelength. The shape of the conductor is considered. The derived formulas show, in particular, the limiting case of infinitesimal proximity of conductors (wherein the grating becomes a corrugated surface).

The fields in the vicinity of the conductors are determined. A system of equivalent boundary conditions is derived for the grating. The Lamb error in the acoustic problem for a grid of circular rods is discussed. Expressions for the fields in the vicinity of the conductors permit taking into account losses caused by nonideal conductors.

1. STATEMENT OF THE PROBLEM. INTRODUCTION

At a plane infinite grid in free space ($\epsilon = \mu = 1$) a plane electromagnetic wave is incident at an oblique angle. The grid is formed by parallel conductors with two axes of symmetry in the cross section. Assuming the period small in comparison with the wavelength it is necessary to determine the transmitted and reflected fields, the field in the immediate vicinity of the conductors, and also to derive local boundary conditions equivalent to the effect of the grid. Solution of the problem is undertaken for investigation of attenuation and phase characteristics of electromagnetic waves in periodic waveguides.

Let us introduce a rectangular system of coordinates x, ξ, y related to the grating as follows (Figure 1). Axis ox is perpendicular to the conductors and lies in the grating plane; 0ξ is directed along the conductors and passes through the center of one of them; oy coincides with the normal to the grating plane. The grating consists of parallel, ideally conducting conductors of arbitrary shape of a cross section having, however, two axes of symmetry, ox and oy. The grating period p is small in comparison with the wavelength λ, so that the dimensionless parameter kp ($k = 2k/\lambda$) is small. Dependence of the fields on time is assumed in the form $\exp(i\omega t)$. From the direction of negative values of the y'-axis (Figure 1), the direction of which in the system x, ξ, y is fixed by direction cosines α, γ, β ($\alpha^2 + \gamma^2 + \beta^2 = 1$), a plane electromagnetic wave of unit amplitude falls upon the grating. The wave reflected from the grating leaves in the negative direction of the oy''-axis, the direction cosines of which are $-\alpha$, $-\gamma$, β. Since the grating is uniform in the ξ direction, dependence of the fields on ξ will be the same as in the incident field, that is, it will be determined by the factor

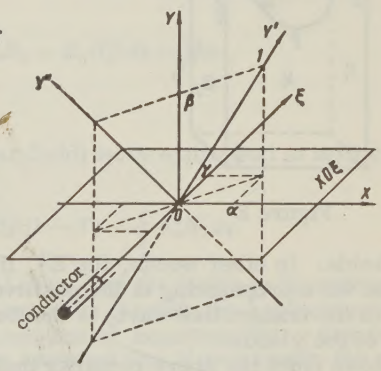


Figure 1

$$e^{-ik\gamma\xi}, \quad \frac{\partial \ldots}{\partial \xi} = -ik\gamma. \tag{1}$$

In the present problem it is necessary to distinguish the two polarizations of the incident wave according to the orientation of its electric and magnetic fields relative to the direction of uniformity of the grating. In the case of H-polarization the electric field is located in the plane perpendicular to the conductors and the magnetic field has a component directed

along the conductors. In the case of E-polarization the magnetic field is perpendicular to conductors and the electric field has a component directed along the conductors. For analysis of the general case it is sufficient to examine the two polarizations individually.

2. CASE OF H-POLARIZATION ($E\xi \equiv 0$)

In this case, in accordance with the Maxwell equations, all components of the field are expressed in terms of $H\xi$ as follows:

$$E_x = \frac{-1}{ik(1-\gamma^2)} \frac{\partial H_\xi}{\partial y}, \quad H_x = \gamma E_y;$$
$$E_y = \frac{1}{ik(1-\gamma^2)} \frac{\partial H_\xi}{\partial x}, \quad H_y = -\gamma E_x. \tag{2}$$

The $H\xi$ component satisfies the Helmholtz equation

$$\nabla^2 H_\xi + k^2(1-\gamma^2) H_\xi = 0 \tag{3}$$

with the boundary condition at the cross-sectional contour of the conductor $\frac{\partial H_\xi}{\partial n}\Big|_c = 0$ (n is the direction of the normal to the cross-sectional contour of the conductor).

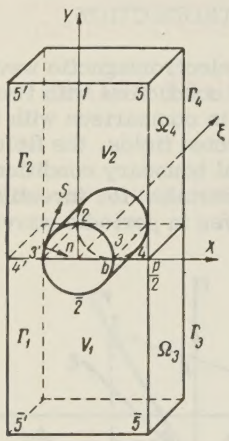


Figure 2

(1) Expression for Fields in Remote Zones in Terms of Fields at the Conductor Contour and at Plane y = 0. Figure 2 shows a typical field sector bounded by volumes V_1 and V_2 on different sides of the grating. Sectors 5'-5 and $\bar{5}'$ -$\bar{5}$ are located at distances from plane y = 0 which ensure formation of the plane fields. In order to determine the relationship between the fields in the remote zones and the fields in the vicinity of the conductors let us apply the Lorentz lemma, (see, for example, reference [1])

$$\oint \{[EH^i]_n - [E^iH]_n\} ds = 0 \tag{4}$$

(n is the direction of the outer normal) to the surfaces bounding volumes V_1 and V_2. As the auxiliary fields figuring in the lemma let us take the fields of the plane waves in free space, choosing their direction of propagation so that the auxiliary waves are "opposed" to the true fields. In other words, the E^1, H^1 field (opposed to the reflected wave) is the field of the wave propagating in the positive direction of the $\bar{y}$''-axis; the E^2, H^2 field (opposed to the transmitted wave) is the field of the wave propagating in the negative direction of the y' axis.

It follows from the above remarks that the quantities $[EH^i]_n$ and $[E^iH]_n$ do not depend on ξ; hence the sum of the integrals over areas lying in planes $\xi = 0$ and $\xi = \xi_0$ yields zero. As a result we obtain zero for the integrals evaluated over the Γ_i contours in the plane $\xi = 0$.*

In vertical sectors $\bar{5}'$ - $\bar{5}'$ and $\bar{5}$ - $\bar{5}$ the following relationships apply:

$$H\left(\frac{p}{2}\right) = H\left(-\frac{p}{2}\right)e^{-ikap}, \quad E\left(\frac{p}{2}\right) = E\left(-\frac{p}{2}\right)e^{-ikap}, \quad E^i\left(\frac{p}{2}\right) = E^i\left(-\frac{p}{2}\right)e^{ikap},$$
$$H^i\left(\frac{p}{2}\right) = H^i\left(-\frac{p}{2}\right)e^{ikap}. \tag{5}$$

*Instead of the Lorentz lemma we may use Green's theorem in "pure form" and operate with potentials H_ξ and H^i_ξ. However, there is little essential difference between the two methods.

This leads to the vanishing of the integral over the vertical sectors. Considering that at the contour of the conductor the tangential component of electric field E_s is equal to zero, we arrive at the desired relationships linking the far fields with the fields at the conductor contour and at the line y = 0 between conductors:

$$\int_{(5'-5)} (E_x^t H_\xi - E_x H_\xi^t)\,dx = \int_{(3'-2-3)} E_s^t H_\xi\,ds + \int_{(4'-3'),(3-4)} (E_x^t H_\xi - E_x H_\xi^t)\,dx,$$
$$\int_{(5'-5)} (E_x H_\xi^t - E_x^t H_\xi)\,dx = \int_{(3-2-3')} E_s^t H_\xi\,ds - \int_{(4'-3'),(3-4)} (E_x^t H_\xi - E_x H_\xi^t)\,dx. \tag{6}$$

The true far fields and the auxiliary fields are:

$$\text{for } y<0 \quad H_{\xi 1} = \sqrt{1-\gamma^2}\,(e^{-ikv'} + Re^{ikv''}),$$
$$\text{for } y>0 \quad H_{\xi 2} = \sqrt{1-\gamma^2}\,Te^{-ikv'}, \tag{7}$$
$$H_\xi^1 = \sqrt{1-\gamma^2}\,e^{-ikv''},\ H_\xi^2 = \sqrt{1-\gamma^2}\,e^{ikv'},$$

R and T are the sought reflection and transmission coefficients. Inserting (7) into (6), we find

$$\int_{(3'-2-3)} E_s^1 H_\xi\,ds + \int_{(4'-3'),(3-4)} (E_x^1 H_\xi - E_x H_\xi^1)\,dx = 0,$$
$$\int_{(3-2-3')} E_s^1 H_\xi\,ds - \int_{(4'-3'),(3-4)} (E_x^1 H_\xi - E_x H_\xi^1)\,dx = -2\beta pR,$$
$$\int_{(3'-2-3)} E_s^2 H_\xi\,ds + \int_{(4'-3')(3-4)} (E_x^2 H_\xi - E_x H_\xi^2)\,dx = -2\beta pT, \tag{8}$$
$$\int_{(3-2-3')} E_s^2 H_\xi\,ds - \int_{(4'-3')(3-4)} (E_x^2 H_\xi - E_x H_\xi^2)\,dx = 2\beta p.$$

Thus, coefficients R and T (that is, of the far field) were expressed in terms of the fields at the contour of the conductor c:

$$-2\beta pR = \oint_c E_s^1 H_\xi\,ds, \qquad 2\beta p\,(1-T) = \oint_c E_s^2 H_\xi\,ds. \tag{9}$$

(2) Fields in the Vicinity of the Conductors. Let us find the fields in the vicinity of the conductors of the grating. Replacement of wave equation (3) with the Laplace equation leads to an error of the order of $(kp)^2$. Hence, if the solution of equation $\nabla^2 H = 0$ is found with an accuracy of the order of $(kp)^2$, it may then be assumed that it represents the true field with the same accuracy. As seen from (8), this solution will be used by us only at the contour of the conductor and at line y = 0 between conductors. Thus, we shall seek a quasi-static solution satisfying the obvious requirements:

(1) the quantities $H\xi$ and $\partial H\xi/\partial y$ are continuous at 4' - 3' and at 3 - 4;

(2) $\partial H\xi/\partial n = 0$ at the cross-sectional contour of the conductor.

In order to determine the form of the solution let us carry out conformal mappings of $z_1(z)$ and $z_2(z)$ translating regions D_1 and D_2 of the planes of complex variables $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ into region D of complex variable $z = x + iy$. Correspondence of the points is shown in Figure 3. Solution of the problem in

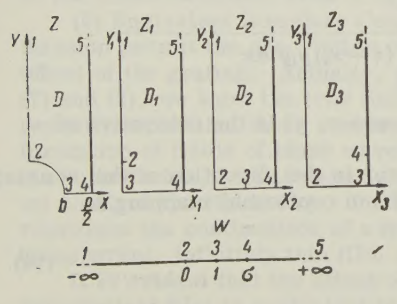


Figure 3.

the entire region under discussion will be sought in the form

$$\text{for } y \geqslant 0 \quad H_{\xi 2} = (A_2 + B_2 y_1)(1 + C_2 x_2),$$
$$\text{for } y \leqslant 0 \quad H_{\xi 1} = (A_1 + B_1 y_1)(1 + C_1 x_2). \tag{10}$$

Here $y_1(x,y)$ and $x_2(x,y)$ are the imaginary and real parts of the transformations of z_1 and z_2; A_k, B_k and C_k are constant quantities. Let us emphasize that at sector 2-3 $\partial y_1 / \partial n \sim \partial x_2 / \partial n = 0$. Let us determine functions y_1 and x_2 over the boundaries of region D from the conditions

$$y_1(-x, y) = y_1(x, y), \; y_1(x, -y) = -y_1(x, y),$$
$$x_2(-x, y) = -x_2(x, y), \; x_2(x, -y) = x_2(x, y). \tag{11}$$

Solution will be subject to the stated conditions.

(1) in sector 3-4 $H_{\xi 2} = H_{\xi 1}$ and $y_1 = 0$. This gives $A_1 - A_2 = (C_2 A_2 - C_1 A_1) x_2$ and, since $x_2 = \text{const}$, this equality is possible only with $A_1 = A_2$, $C_1 = C_2$ (let us designate $A_1 = A$, $C_1 = C$);

(2) from the continuity of E_x in sector 3-4 it follows that $B_1 = B_2$ (let us designate $B_1 = B$);

(3) finally, the conditions in (5) permit easy determination of $C = -ik\alpha$. As a result, a field in the vicinity of the conductors may be represented in terms of two constants A and B in the form

$$H_\xi = (A + B y_1)(1 - ik\alpha x_2). \tag{12}$$

Quantities A and B are determined from the first and fourth equations in (8). It is shown below that B is of the order of kp; hence $\nabla^2 H$ is of the order of $(kp)^2$, that is, the desired solution represents the true field with the assumed accuracy.

(3) Calculation of R, T, A and B. Now, by use of relationships (8) and (9) and the derived fields in the vicinity of the conductors, it is not difficult to calculate the quantities which we seek. Let the cross-sectional contour of the conductor be given by the equation $y = y(x)$. In evaluating the integrals along the conductor contour let us replace the exponents by the first two terms of expansion for ky (the vertical dimensions of the conductor are assumed to be small in comparison with the wavelength).

As a result of integration we find

$$A = \frac{\beta \sqrt{1-\gamma^2}}{\beta - ik[(1-\gamma^2)l + \alpha^2 L]}, \quad B = \frac{-ik\sqrt{1-\gamma^2}\beta}{1 + ikp\Delta},$$
$$R = \frac{1}{2}\left\{\frac{\beta + ik[(1-\gamma^2)l + \alpha^2 L]}{\beta - ik[(1-\gamma^2)l + \alpha^2 L]} - \frac{1 - ikp\bar{\Delta}}{1 + ikp\Delta}\right\}, \tag{13}$$
$$T = \frac{1}{2}\left\{\frac{\beta + ik[(1-\gamma^2)l + \alpha^2 L]}{\beta - ik[(1-\gamma^2)l + \alpha^2 L]} + \frac{1 - ikp\bar{\Delta}}{1 + ikp\Delta}\right\}.$$

Here $l = S/2p$ (S is the cross-sectional area of the conductor);

$$\Delta = \beta\Delta_1 - ikp(\beta^2\Delta_2 - \alpha^2\Delta_3);$$
$$\Delta_1 = \frac{2}{p^2}\int_0^b y_1\,dx; \; \Delta_2 = \frac{2}{p^3}\int_0^b y y_1\,dx; \; \Delta_3 = \frac{2}{p^3}\int_0^b (x - x_2) y_1 y'\,dx;$$

$L = \frac{2}{p}\int_0^b x_2 y'\,dx$ (the integrands are applied to points of the contour, y' is the derivative of function y(x), 2b is the maximum dimension of the conductor in the direction of the ox axis). It may be shown that $\Delta_1 = l_1/p$, where l_1 is determined from conformal mapping as the limit:

$$l_1 = \lim_{y \to \infty}(y_1 - y). \tag{14}$$

(4) Checking the Limiting Cases. It is evident that the obtained coefficients of reflection

and transmission satisfy the law of conservation of energy $R\bar{R} + T\bar{T} = 1$. With $q = 2b/p \to 0$ ($b \to 0$) $T \to 1$, $R \to 0$. The same transition to the limit occurs with $kp \to 0$ (in other words, a decrease in period leads to an increase in penetration of the field through the grating).

Let us explain the behavior of R and T with $q \to 1$ (transition to solid filling, that is, to corrugation). For this purpose let us make still another transformation; that is, let us carry out conformal representation of the upper semiplane of the auxiliary complex plane W (Figure 3) in the region D_1 of plane z_1 so that points $-\infty, 0, 1, \sigma, +\infty$ of the real axis W are transferred to points 1,2,3,4,5, respectively, of plane z_1. The transformation is written in the form

$$z_1 = \frac{p}{\pi} \arcsin \sqrt{\frac{W-1}{\sigma-1}}. \tag{15}$$

Parameter σ depends on the shape and relative dimensions of the conductor. For calculation of Δ it is necessary to know the values of y_1 at sector 2-3. At this sector $0 \leqslant W \leqslant 1$, hence,

$$y_1 = J_m z_1 = \frac{p}{\pi} \ln \frac{\sqrt{1-W} + \sqrt{\sigma-W}}{\sqrt{\sigma-1}}. \tag{16}$$

With $q \to 1$ (the conductors approach one another infinitesimally $b \to p/2$) $\sigma \to 1$, and $y_1 \to \infty$ (let it be remembered that it was not required that $ky_1 \ll 1$, and only that $ky \ll 1$). Then, from the integrals defining Δ the principal terms tending toward infinity are isolated. (It may be shown that the singularity of expression in $(\sqrt{1-W} + \sqrt{\sigma-W})$ at the point $W = 1$ for $\sigma = 1$ is integrable);

$$\begin{gathered} \Delta_1 \simeq \frac{y_1}{p}, \quad \Delta_2 \simeq \frac{y_1}{p} l, \quad \Delta_3 \simeq -\frac{y_1}{p}(l+L), \\ \beta^2\Delta_2 - \alpha^2\Delta_3 \simeq \frac{y_1}{p}[(1-\gamma^2)l + \alpha^2 L]. \end{gathered} \tag{17}$$

Inserting this into the formula for R and T, we find for $q = 1$

$$R = \frac{\beta + ik[(1-\gamma^2)l + \alpha^2 L]}{\beta - ik[(1-\gamma^2)l + \alpha^2 L]}. \tag{18}$$

In addition, $T = 0$, $|R| = 1$, which was also to be expected. With normal incidence $\alpha = \gamma = 0$. $\beta = 1$. Formula (18) for R in our approximation may be written in the form

$$R = e^{ik\frac{S}{p}}. \tag{19}$$

For a rectangular waveguide this formula, as it should, yields a phase excursion with simple displacement of an ideal conducting plane in the negative direction of y. Thus, with normal incidence of a plane wave of H-polarization on a corrugated plane surface with a corrugation of arbitrary shape, the phase excursion in the reflection coefficient is proportional to the cross-sectional area of the figure forming the corrugation.

(5) Equivalent Boundary Conditions. The derived coefficients of reflection and transmission permit the description of local boundary conditions which are equivalent to the effect of the grating. Actually, after calculation of R and T, in accordance with formulas (7) and (2), we know the true fields for regions lying above line 5'-5 and below line $\bar{5}' - \bar{5}$ (we will recall that these sectors are located from plany $y = 0$ at distances sufficient for the formation of fields of plane waves. Usually, these distances are of the order of a period, since the fields of local waves attenuate rapidly with distance from the grating.) If now we continue the fields to $y = 0$ as fields of plane waves (remembering that this does not represent the continuation of a true field), we thereby introduce with $y = 0$ a certain semitransparent, infinitely thin film at which the jumps in the components of the field are defined.

It is evident that the effect of this hypothetical film is equivalent to the effect of this hypothetical film is equivalent to the effect of a real grating, since the fields in the remote zones are preserved. The jumps in field components may be regarded as the equivalent

boundary conditions which must be laid down with y = 0. In this manner we obtain the following boundary conditions for waves of H-polarization:

$$H_{\xi 2} - H_{\xi 1} = \frac{-ikl_1}{1+(kp)^2(\beta^2\Delta_2 - \alpha^2\Lambda_3)}\left[E_{x2} + E_{x1} + \frac{i}{k}\frac{\partial}{\partial\xi}(H_{y2} + H_{y1})\right],$$
$$E_{x2} - E_{x1} = ik\left[l(H_{\xi 2} + H_{\xi 1}) - \frac{i}{k}\frac{\partial}{\partial\xi}(E_{y2} + E_{y1})\right]. \tag{20}$$

Consideration of the component of the order of $(kp)^2$ in the denominator of the first boundary condition, as in the formulas for R and T, makes sense for q sufficiently close to unity when parameters Δ_2 and Δ_3 begin to increase substantially. These results are a generalization of the results given in reference [2] (for $\epsilon = 1$) for the case where the fields are dependent on coordinate ξ ($\partial \ldots /\partial\xi \neq 0$).

3. CASE OF E-POLARIZATION ($H\xi \equiv 0$)

In this case all the components of the field may be expressed in terms of the component of the electric field along the conductor as follows:

$$H_x = \frac{1}{ik(1-\gamma^2)}\frac{\partial E_\xi}{\partial y}, \quad E_x = -\gamma H_y,$$
$$H_y = \frac{-1}{ik(1-\gamma^2)}\frac{\partial E_\xi}{\partial x}, \quad E_y = \gamma H_x. \tag{21}$$

The same component E_ξ satisfies the equation $E_\xi + k^2(1-\gamma^2)E_\xi = 0$ with the boundary condition at the conductor contour $E_\xi|_c = 0$.

(1) Expression for Far Fields in Terms of the Field at the Contour and at the y =0 Plane. Application of the Lorentz lemma to volumes V_1 and V_2 leads here to the relationships:

$$\int_{(5'-5)} (E^1_\xi H_x - E_\xi H^1_x)\,dx = \int_{(3'-2-3)} E^1_\xi H_s\,ds + \int_{(4'-3'),\,(3-4)} (E^i_\xi H_x - E_\xi H^i_x)\,dx,$$
$$\int_{(5'-5)} (E_\xi H^1_x - E^1_\xi H_x)\,dx = \int_{(3-2-3')} E^1_\xi H_s\,ds - \int_{(4'-3'),\,(3-4)} (E^i_\xi H_x - E_\xi H^i_x)\,dx. \tag{22}$$

The true far fileds and the auxiliary fields are

$$\text{for } y>0 \quad E_{\xi 2} = T\sqrt{1-\gamma^2}\,e^{-ik\gamma y'},$$
$$\text{for } y<0 \quad E_{\xi 1} = \sqrt{1-\gamma^2}\,(e^{-ik\gamma y'} + Re^{ik\gamma y''}),$$
$$E^1_\xi = \sqrt{1-\gamma^2}e^{-ik\gamma y''}, \quad E^2_\xi = \sqrt{1-\gamma^2}\,e^{ik\gamma y'}. \tag{23}$$

Inserting (23) into (22), let us find the relation of the desired coefficients R and T to the fields at the conductor contour:

$$\int_{(3'-2-3)} E^1_\xi H_s\,ds + \int_{(4'-3'),\,(3-4)} (E^1_\xi H_x - E_\xi H^1_x)\,dx = 0,$$
$$\int_{(3-2-3')} E^1_\xi H_s\,ds - \int_{(4'-3'),\,(3-4)} (E^1_\xi H_x - E_\xi H^1_x)\,dx = 2\beta pR,$$
$$\int_{(3'-2-3)} E^2_\xi H_s\,ds + \int_{(4'-3'),\,(3-4)} (E^2_\xi H_x - E_\xi H^2_x)\,dx = -2\beta pT, \tag{24}$$
$$\int_{(3-2-3')} E^2_\xi H_s\,ds - \int_{(4'-3'),\,(3-4)} (E^2_\xi H_x - E_\xi H^2_x)\,dx = 2\beta p,$$

$$-2\beta pR = \oint_c E_\xi^1 H_s\, ds,$$
$$2\beta p(1-T) = \oint_c E_\xi^2 H_s\, ds. \tag{25}$$

(2) Fields in the Vicinity of Conductors. Let us seek the fields in the vicinity of the conductors from the solution of the electrostatic problem, which may be formulated in the following manner. It is necessary to find the solution of E_ξ of the Laplace equation (with an accuracy of the order of $(kp)^2$) satisfying the requirements:

(1) the quantities E_ξ and $\partial E_\xi / \partial y$ are continuous at 4'-3', 3-4;

(2) $E_\xi = 0$ at the cross-sectional contour of the conductor. A similar electrostatic problem is solved in reference [3] for the case of independence of the fields on coordinate ξ.

In order to determine the form of the solution let us carry out conformal representation of $z_3(z)$, translating region D_3 of plane $z_3 = x_3 + iy_3$ to region D of complex plane $z = x + iy$ (Figure 3). Solution of the problem has the form

$$\text{for } y \geqslant 0 \quad E_{\xi 2} = (A_2 y_2 + B_2 y_3)(1 + C_2 x_2),$$
$$\text{for } y \leqslant 0 \quad E_{\xi 1} = (A_1 y_2 + B_1 y_3)(1 + C_1 x_2). \tag{26}$$

Here $y_2(x,y)$ and $y_3(xy)$ are the imaginary parts of transformations of $z_2(z)$ and $z_3(z)$; $x_2(x,y)$ is the real part of transformation of $z_2(z)$; A_n, B_n and C_n are constants. Let us determine the functions of y_2 and y_3 beyond the limits of region D from the conditions

$$y_2(-x,y) = y_2(x,y),\ y_2(x,-y) = -y_2(x,y),$$
$$y_3(-x,y) = y_3(x,y),\ y_3(x,-y) = y_3(x,y). \tag{27}$$

As in the case of H-polarization, one of the functions is extended beyond y = 0 symmetrically and the other asymmetrically.

(1) At sector 3-4 $E_{\xi 2} = E_{\xi 1}$. This gives $B_1 - B_2 = (B_1C_1 - B_2C_2)x_2$, whence $B_1 = B_2$, $C_1 = C2$ (let us designate $B_1 = B$, $C_1 = C$).

(2) From the continuity of H_x ($H_x = \frac{1}{ik(1-\gamma^2)} \frac{\partial E_\xi}{\partial y}$) at 3-4 it follows that $A_1 = A_2$ (let us designate $A_1 = A$).

(3) The conditions in (5), as before, give $C = ik\alpha$. Thus, the sought field in the vicinity of the conductors may be written in the form

$$E_\xi = (Ay_2 + By_3)(1 - ik\alpha x_2). \tag{28}$$

Constants A and B are determined from the first and fourth equations of (24).

(3) Calculation of A, B, R, and T. For calculation of A and B let us add and subtract the first and fourth equations of (24). We obtain

$$\int_{(3'-2-3)} E_\xi^1 H_s\, ds + \int_{(3-2-3')} E_\xi^2 H_s\, ds - 2 \int_{(4'-3'),\ (3-4)} H_x^1 E_\xi\, dx = 2\beta p,$$
$$-\int_{(3'-2-3)} E_\xi^1 H_s\, ds + \int_{(3-2-3')} E_\xi^2 H_s\, ds - 2 \int_{(4'-3'),\ (3-4)} E_\xi^1 H_x\, dx = 2\beta p. \tag{29}$$

At sectors 4'-3' and 3-4, $y_2 = 0$ and E_ξ may be represented in the form

$$E_\xi = By_3 e^{-ik\alpha x_1},\ H_x^1 = \frac{-\beta}{\sqrt{1-\gamma^2}} e^{ik\alpha x}.$$

Hence,

$$\int_{(4'-3'),\ (3-4)} H_x^1 E_\xi\, dx = \frac{-\beta B}{\sqrt{1-\gamma^2}} \int_{(4'-3'),\ (3-4)} y_3\, dx.$$

At this same sector $\partial x_2 / \partial y = 0$ and $\partial y_3 / \partial y = 0$, whence

$$H_x = \frac{A}{ik(1-\gamma^2)} \frac{\partial y_2}{\partial y}$$

and

$$\int\limits_{(4-3'),\,(3-4)} E_\xi^1 H_x \, dx = \frac{A}{ik(1-\gamma^2)} \int\limits_{(4'-3'),\,(3-4)} \frac{\partial y_2}{\partial y} dx.$$

It is required that we evaluate integrals along the conductor contour of the form $\int E_\xi^i H_s \, ds$, but $H_s = \frac{1}{ik(1-\gamma^2)} \frac{\partial E\xi}{\partial n}$ and, since at the conductor contour $y_2 = y_3 = 0$,

$$\frac{\partial E_\xi}{\partial n} = \left(A \frac{\partial y_2}{\partial n} + B \frac{\partial y_3}{\partial n} \right) e^{-ik\alpha x},$$

$$\int E_\xi^i H_s \, ds = \frac{-1}{ik(1-\gamma^2)} \int E_\xi^i \left(A \frac{\partial y_2}{\partial n} + B \frac{\partial y_3}{\partial n} \right) e^{-ik\alpha x} dx.$$

Inserting these expressions in formula (29), we find

$$\begin{aligned} B\left[\int\limits_{(3'-2-3)} \frac{\partial y_3}{\partial n} ds - ik\beta \left(\int\limits_{(3'-2-3)} y \frac{\partial y_3}{\partial n} ds + \int\limits_{(4'-3'),\,(3-4)} y_3 dx \right) \right] &= -ik\sqrt{1-\gamma^2}\,\beta p, \\ A\left[\int\limits_{(3'-2-3)} \frac{\partial y_2}{\partial n} ds - \int\limits_{(4'-3'),\,(3-4)} \frac{\partial y_2}{\partial y} dx - ik\beta \int\limits_{(3'-2-3)} y \frac{\partial y_2}{\partial n} ds \right] &= ik\sqrt{1-\gamma^2}\,\beta p. \end{aligned} \tag{30}$$

Let us evaluate the integrals in the left-hand sides of the equations by means of Green's formula. We write the condition of equality to zero for the integrals over the closed contour Γ_2 from the functions $\partial y_2 / \partial n$, $\partial y_3 / \partial n$, $y\frac{\partial y_2}{\partial n}$, $-y_2\frac{\partial y_3}{\partial n}$, $y\frac{\partial y_3}{\partial n}$, $-y_3\frac{\partial y}{\partial n}$.

For the mentioned conformal representations there occurs the limiting relationship $dz_i / dz \to 1$ (i = 1, 2, 3) with $y \to \infty$; hence $y_i = y + l_i$. From this circumstance we find

$$\begin{aligned} &\int\limits_{(3'-2-3)} \frac{\partial y_3}{\partial n} ds = -p, \quad \int\limits_{(3'-2-3)} \frac{\partial y_3}{\partial n} ds - \int\limits_{(4'-3'),\,(3-4)} \frac{\partial y_2}{\partial y} dx = -p, \\ &\int\limits_{(3'-2-3)} y \frac{\partial y_2}{\partial n} ds = pl_2, \quad \int\limits_{(3'-2-3)} y \frac{\partial y_3}{\partial n} ds + \int\limits_{(4'-3'),\,(3-4)} y_3 dx = pl_3. \end{aligned} \tag{31}$$

Inserting (31) into (30), we obtain A and B (i.e., the field in the vicinity of the conductors):

$$A = \frac{-ik\beta\sqrt{1-\gamma^2}}{1+ik\beta l_2}, \quad B = \frac{ik\beta\sqrt{1-\gamma^2}}{1+ik\beta l_3}. \tag{32}$$

Now from formulas (25) we may obtain the desired coefficients of reflection and transmission:

$$\begin{aligned} R &= -\frac{1}{2}\left[1 - 2ik\beta l_2 + \frac{1-ik\beta l_3}{1+ik\beta l_3} \right], \\ T &= \frac{1}{2}\left[1 - 2ik\beta l_2 - \frac{1-ik\beta l_3}{1+ik\beta l_3} \right]. \end{aligned} \tag{33}$$

Parameters l_i depend on the shape and relative dimensions of the conductor. Formulas (33) permit, in particular, both limit transitions ($q \to 0$, $q \to 1$). It is further shown that for $q \to 0$, $l_2 \to 0$ and $l_3 \to \infty$, hence $T \to 1$ and $R \to 0$; for $q \to 1$, $l_2 \to l_3$, hence $T \to 0$ and $R \to e^{-2ik\beta l_3}$.

(4) Equivalent Boundary Conditions. The derived expressions for the coefficients of reflection and transmission, as in the case of H-polarization, permit writing the boundary conditions equivalent to the effect of a real grating in the following form:

$$\begin{aligned} E_{\xi 2} - E_{\xi 1} &= ikl_2 \left[(H_{x2} + H_{x1}) - \frac{i}{k} \frac{\partial}{\partial \xi} (E_{y2} + E_{y1}) \right], \\ E_{\xi 2} + E_{\xi 1} &= ikl_3 \left[(H_{x2} - H_{x1}) - \frac{i}{k} \frac{\partial}{\partial \xi} (E_{y2} - E_{y1}) \right]. \end{aligned} \tag{34}$$

For the special case of a grating of ribbon conductors these conditions coincide with the boundary conditions obtained in reference [4]. Boundary conditions (20) and (34) may be used for analysis of attenuation and the phase characteristics of electromagnetic waves in helical waveguides taking into account the period, shape, and relative dimensions of the conductor for the entire range of changes in parameter 2 (from a grating of extremely fine conductors to a corrugated surface).

4. CALCULATION OF PARAMETERS. CONDUCTORS OF CIRCULAR AND RECTANGULAR CROSS SECTION

In order to obtain the more convenient formulas defining the l_i parameters let us carry out conformal representations of $z_i(W)$ (i = 1, 2, 3) translating regions D_i of the planes of complex variables of z_i to the upper semiplane of complex variable W (Figure 3). These transformations may be presented in the form

$$z_1 = \frac{p}{\pi} \arcsin \sqrt{\frac{W-1}{\sigma-1}},$$
$$z_2 = \frac{p}{\pi} \arcsin \sqrt{\frac{W}{\sigma}}, \tag{35}$$
$$z_3 = \frac{p}{\pi} \arcsin \sqrt{W}.$$

The correspondence of points is shown in Figure 3. Formulas for calculating l_2 and l_3 may be presented in the form

$$l_i = l_1 + \lim_{\substack{y \to \infty \\ (W \to \infty)}} (y_i - y_1). \tag{36}$$

Using the transformations of (35) for calculation of the limits, we find

$$l_2 = l_1 - \frac{p}{2\pi} \ln \frac{\sigma - 1}{\sigma}.$$
$$l_3 = l_1 + \frac{p}{2\pi} \ln (\sigma - 1). \tag{37}$$

The formulas are valid for conductors of any shape. By means of the Cauchy-Riemann conditions it may be shown that parameter L in the formulas of H-polarization is equal to l_2.

(1) Conductor of a Circular Cross Section. For a grid of circular conductors we may use Blokh's extremely precise transformation [5].

$$z_1 = z - \frac{mb}{2\pi} \ln \frac{\sin \frac{\pi}{p}(z + \lambda b)}{\sin \frac{\pi}{p}(z - \lambda b)}, \tag{38}$$

where b is the conductor radius, m and λ are parameters depending of q. For parameter l_1 the following formula is obtained:

$$l_1 = \frac{m\lambda}{4} pq^2.$$

Using the transformation of $z_1(W)$, we find

$$\sigma = \operatorname{cth}^2 r,$$
$$r = \frac{\pi q}{2}\left[1 + \frac{m}{\pi} \operatorname{arc\,tg}\left(\operatorname{tg} \frac{\pi}{2} q\lambda \operatorname{cth} \frac{\pi q}{2}\right)\right]$$

(here $r \frac{p}{\pi}$ is the ordinate of point 2 at plane z_1), whence $l_2 = l_1 - \frac{p}{\pi}$ ln cosh r and $l_3 = l_1 - \frac{p}{\pi}$ in sinh r.

for $q \to 0$ $r \to 0$, $\sigma \to \infty$, $l_1 \to \infty$, $l_2 \to 0$, $l_3 \to \infty$.
for $q \to 1$ $\lambda \to 1$, $r \to \infty$, $l_2 \to l_3 \to 0{,}45$.

In the case of H-polarization, normal incidence and a fine conductor (q 1) use of Blokh's transformation leads to formulas

$$R = 3ikl, \quad T = 1 - ikl, \tag{39}$$

coinciding with the formulas derived in [6]. Lamb's results [7], derived for this case in the

acoustic problem of a grating of thin circular rods with the identical statement of the problem, and Gans' results [8] are erroneous. Their formulas did not consider the annular current along the perimeter of the conductor, which led to an error of 1.5 times for the reflection coefficient.

(2) Conductor of Rectangular Cross-Section. Ribbon Conductor. For a conductor of rectangular cross section, by means of the Christoffel-Schwartz formula we may obtain the transformation of z(W) in the form

$$z = \frac{p}{2\pi} \int_0^W \frac{1}{\sqrt{U(1-U)}} \sqrt{\frac{U-t}{U-\sigma}}\, dU + ic. \tag{40}$$

Here 2c is the conductor dimension in the direction of the oy axis. Parameters σ and t, depending on the relative dimensions of the conductor, are determined from Equations (5) and (6) in [9]. Calculation of σ and t was performed for several values of relative dimensions in this work. From transformations of z_1 (W) and z(W) from formula $l_1 = \lim_{W \to \infty} (y_1 - y)$ we find

$$l_1 = \frac{pt}{2\pi} \int_0^{\frac{1}{\sigma}} \frac{dU}{\sqrt{(1-U)(1-\sigma U)}\left(\sqrt{1-tU}+1\right)}. \tag{41}$$

For a ribbon conductor (c = 0) planes z and z_2 coincide,

$$l_2 = 0,\ \sigma = \frac{1}{\sin^2 \frac{\pi}{2} q},\ l_1 = \frac{p}{\pi} \ln \frac{1}{\cos \frac{\pi}{2} q},\ l_3 = \frac{p}{\pi} \ln \frac{1}{\sin \frac{\pi}{2} q}.$$

5. CALCULATION OF THE COEFFICIENTS OF REFLECTION AND TRANSMISSION BY MEANS OF INTEGRAL FORMS OF MAXWELL EQUATIONS

Solution of the problem in the case of normal incidence of the wave at the grating may be obtained by another, simpler method. For this purpose let us use the Maxwell equation in integral form.

(1) H-Polarization. Here we use the equations

$$\oint_{(\Gamma_1+\Gamma_2)} E_s\, ds = -ik \iint H_\xi\, ds, \tag{42}$$

$$\oint_{(\Gamma_3+\Gamma_4)} H_s\, ds = ik \iint E_x\, ds. \tag{43}$$

For evaluation of the area integrals let us use the derived static solution: $H_\xi = A + By_1$, $E_x = \frac{-B}{ik} \frac{\partial y_1}{\partial y}$. At the verticals $\partial \ldots / \partial n = 0$. Directly from Equation (42) we obtain

$$E_{x2} - E_{x1} = -ikA2y + ik\frac{S}{p} A. \tag{44}$$

The term -ik2y represents the simple phase excursion between sectors $\overline{5}'$ - $\overline{5}$ and 5' - 5. The second term, proportional to the area occupied by the grating, is not considered in the Lamb formulas. Equation 43 gives

$$H_{\xi 2} - H_{\xi 1} = -B2y + B2l_1. \tag{45}$$

Analyzing the solution for far zones in terms of ky and equating the coefficients, we find

$$\left.\begin{matrix} R \\ T \end{matrix}\right\} = \frac{1}{2}\left(1 + 2ikl \mp \frac{1 - ikl_1}{1 + ikl_1}\right), \tag{46}$$

$$A = 1 + ikl,\quad B = \frac{-ik}{1 + ikl_1},\quad l = \frac{S}{2p}.$$

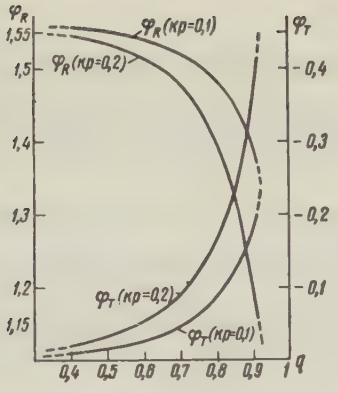


Figure 4

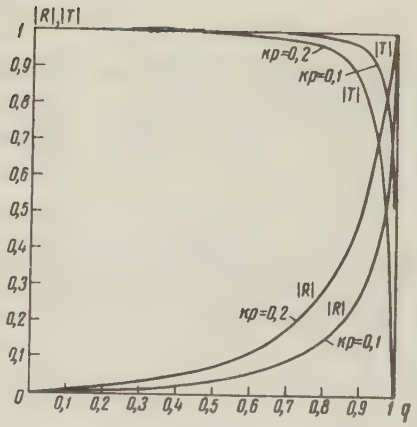


Figure 5

The boundary conditions corresponding to these coefficients will have the form

$$H_{\xi 2} - H_{\xi 1} = -ikl_1(E_{x2} + E_{x1}),$$
$$E_{x2} = E_{x1} = ikl(H_{\xi 2} + H_{\xi 1}). \tag{47}$$

Formulas (46) and (47) are applicable for q not extremely close to unity and do not give the correct critical transition for $q \to 1$. This circumstance is an evident consequence of the use of the static solution at a sufficient distance from the grating.

In the case of H-polarization two types of current are observed at the contour: annular (owing to constant A) and dipole (owing to the term By_1). At the upper half of the contour these points are subtracted and at the lower half they are added. The dipole current is always small as long as the slots are quite small. This explains the excellent transmission of the H-polarized wave through the grating. Figures 4 and 5 show the moduli and phases of R and T as calculated for a grid of circular conductors from the precise formulas (13) for values of parameter kp = 0.1 and 0.2.

(2) E-Polarization. Here we use the equation

$$\oint_{\Gamma_n} E_s\,ds = -ik \iint_{\Omega_n} H_x\,ds \quad (n = 3, 4). \tag{48}$$

Contours Γ_n and areas Ω_n lie in the plane x = p/2. For evaluation of the area integrals let us use the static solution $E_\xi = Ay_2 + By_3$, $H = \frac{A}{ik}\frac{\partial y_2}{\partial y} + \frac{B}{ik}\frac{\partial y_3}{\partial y}$. We obtain

$$E_{\xi 2} - E_{\xi 1} = 2Al_2 + 2Ay,$$
$$E_{\xi 2} + E_{\xi 1} = 2Bl_3 + 2By. \tag{49}$$

Equating the coefficients with powers of y in (49) and in the expressions for the far field, we obtain the formulas

$$\left.\begin{matrix} R \\ T \end{matrix}\right\} = \frac{1}{2}\left(1 - 2ikl_2 \pm \frac{1 - ikl_3}{1 + ikl_3}\right), \tag{50}$$

which are a special case of formulas (32) and (33)

The derived expressions for fields in the vicinity of the conductors permit calculation (from existing formulas) of losses associated with nonideal conductors.

In conclusion I wish to express my thanks of B.Z. Katsenelenbaum for his interest and valuable advice in this work.

REFERENCES

1. L.A. Vainshteyn. Elektromagnitye volny [Electromagnetic waves], Izd. Sovetskoye Radio, 1956, 1. 418.
2. A.N. Sivov, Incidence of a plane electromagnetic wave at a plane grating (case wherein the H vector is parallel to the conductors) Radiotekhnika i elektronika, 1959, 6, 1, 58.
3. B.Z. Katsenelenbaum, Attenuation of H_{0n} waves in a helical waveguide, Radiotekhnika i elektronika, 1959, 4, 3, 428.
4. N.N. Smirnov, Propagation of electromagnetic waves in circular waveguides with periodic slots, ZhTF, 1958, 28, 7, 1494.
5. E. Blokh, Investigation of a plane grating consisting of theoretical profiles of finite thickness, Tr. TsAGI, 1947, No. 611[
6. B.Z. Katsenelenbaum, Determining the coefficient of reflection of an electromagnetic wave in a rectangular waveguide from a thin metal rod perpendicular to the electric field, Izv, vuzov MVO, SSSR (Radio tekhnika), 1960, 5.
7. G. Lamb, Gidrodynamika, [Hydrodynamics], GITTL, p. 673.
8. Von R. Gans, Das verhalten Hertzscher Gitter, Ann. Physic, 1920, 61, 5, 447.
9. V.V. Malin, A.N. Sivov, Toward a theory of propagation of H_{01} waves in a spiral waveguide, Radiotekhnika i elektronika, 1959, 4, 3, 433.

Submitted to the editors 23 July 1960

FLUCTUATIONS IN SELF-EXCITED OSCILLATORS

S.I. Yevtyanov and V.N. Kuleshov

Using the methods of shortened symbolic equations, general expressions are derived describing the effect of a small additive noise current on the amplitude and phase of sinusoidal oscillations of a self-excited oscillator with any number of degrees of freedom. On the basis of these expressions there is constructed a linear equivalent circuit for fluctuation.

The presented equivalent circuit is suitable for investigation of local stability of stationary operating conditions.

INTRODUCTION

It is the purpose of this report to derive general formulas for calculation of fluctuations in self-excited sine-wave oscillators. The general procedure for calculation of fluctuations in such oscillators has been developed in a number of papers [1 - 5]. Nevertheless, as we well know, the literature still contains no work in which expressions for the statistical characteristics of oscillations in a wide class of self-excited oscillators have been derived in closed form.

Below we derive such expressions for a self-excited oscillator containing a vacuum tube and a four-terminal feedback network in the form of a system of low-loss tuned circuits. In this case the behavior of the self-excited oscillator is conveniently investigated by the method of shortened symbolic equations[6]. In reference [7] this method was used for calculating the fluctuations in single-tuned circuits.

In deriving the formulas the influence of plate and grid current response is disregarded. The grid is assumed to be fixed-based or self-biased but lagless.

1. DERIVATION OF EQUATIONS FOR FLUCTUATION

Let $\dot{I}$ and $\dot{U}$ represent the complex amplitudes of the first harmonic of plate current and control (grid) voltage and let φ represent the possible phase shift between self-oscillation and oscillation at a certain reference frequency ω_0:

$$\dot{I} = Ie^{i\varphi}, \tag{1}$$

$$\dot{U} = Ue^{i\varphi}. \tag{2}$$

If frequency ω_0 is chosen within the passband limits of the four-terminal feedback network, then the amplitudes of I and U and phase φ will be slowly varying functions of time. The phases of the first harmonics of current and voltage are assumed to be identical (i.e., electron inertia within the tube is disregarded).

The relation between the amplitudes of the first harmonics of plate current and control voltage is determined by the oscillatory characteristics of the self-excited oscillator.

$$I = I(U). \tag{3}$$

Its form depends on the shape of the static characteristics of the tube, the bias at the grid or the parameters of lagless self-biased meshes.

On the other hand, the complex amplitudes $\dot{I}$ and $\dot{U}$ are related by the shortened symbolic equation

$$y(p)U = \dot{I}. \tag{4}$$

Here y(p) is the symbolic control (grid) admittance. It is obtained after contraction of the expression $y(i\omega_0 + p)$ if it is assumed that $p = i\Omega$, where Ω is a small detuning relative to ω_0 (reference [6]).

Let us now assume that a noise current is added to current $\dot{I}$ in the plate circuit. This noise current may also be characterized by a complex amplitude $\dot{I}_n$, having in mind analysis into two quadrature oscillations with random amplitudes

$$\dot{I}_n = I_a + iI_p. \tag{5}$$

Here I_a and I_p are the real and imaginary components of noise current. To these components we will ascribe the power spectra $\overline{I^2_{2\Omega}}$ and $\overline{I^2_{p\Omega}}$. Taking into account the noise current Equation (4) takes the form

$$y(p)\dot{U} = \dot{I} + \dot{I}_n \tag{6}$$

Solution of this equation is sought, as in [7], with an accuracy of the order of the square of the fluctuations in amplitude and phase in the form

$$\dot{I} = (I + I_f + iI\varphi_f)e^{i\varphi}, \tag{7}$$

$$\dot{U} = (U + U_f + iU\varphi_f)e^{i\varphi}. \tag{8}$$

Here I and U are the stationary amplitudes of the first harmonic of current and voltage of self-oscillation; φ, the stationary difference in phase between self-oscillation and oscillation with frequency ω_0, may be constant or a linear function of time:

$$\varphi = \Delta\omega t + \varphi_0. \tag{9}$$

In the latter case, the difference between the frequency of self-oscillation and frequency ω_0 is Δ_ω (correction for frequency).

In (7) and (8) I_f, U_f and φ_f (as in reference [7]) represent the fluctuational additions to the amplitude and phase of self-oscillation.

In order to determine the stationary conditions let us assume first that $I_f = 0$, $U_f = 0$, $\varphi_f = 0$ and insert (7) and (8) into (4). Considering (9) and using the displacement theorem we obtain the following equation:

$$y(p + i\Delta\omega)U = I. \tag{10}$$

Let us divide $y(p + i\Delta\omega)$ into real and imaginary parts:

$$y(p + i\Delta\omega) = y_{\text{real}}(p + i\Delta\omega) + iy_{\text{imag}}(p + i\Delta\omega) \tag{11}$$

and insert (11) into (10), assuming p = 0. The resulting equations for stationary operation are

$$y_{\text{real}}(i\Delta\omega)U = I, \tag{12}$$

$$y_{\text{imag}}(i\Delta\omega) = 0. \tag{13}$$

Together with Equation (3) they define the values of U, I, and $\Delta\omega$. We may then obtain the fluctuation equations if we insert (5), (7) and (8) into (6) and keep in mind equalities (11) (12) and (13):

$$y_{\text{real}}(p + i\Delta\omega)U_f - I_f - y_{\text{imag}}(p + i\Delta\omega)U\varphi_f = I_a, \tag{14}$$

$$y_{\text{imag}}(p + i\Delta\omega)U_f + y_{\text{real}}(p + i\Delta\omega)U\varphi_f - I\varphi_f = I_p. \tag{15}$$

If we let S represent the mean transconductance of the oscillatory characteristic and σ its local slope at the stationary operating point, then the equation for relation between U and I and between U_f and I_f [7] may be presented in the form

$$I = SU, \tag{16}$$

$$I_f = \sigma U_f. \tag{17}$$

With these expressions, and introducing the designations

$$g = S - \sigma, \tag{18}$$

$$Y_{22}(p) = y_{\text{real}}(p + i\Delta\omega) - S, \tag{19}$$

$$Y_{11}(p) = Y_{22}(p) + g, \tag{20}$$

$$Y_{12}(p) = y_{\text{imag}}(p + i\Delta\omega). \tag{21}$$

we obtain the fluctuation equations in the final form

$$Y_{11}(p)U_f - Y_{12}(p)U\varphi_f = I_a, \tag{22}$$

$$Y_{12}(p)U_f + Y_{22}(p)U\varphi_f = I_p. \tag{23}$$

3. EQUIVALENT CIRCUIT FOR FLUCTUATIONS AND SOLUTION OF EQUATIONS (22) AND (23)

Equations (22) and (23) are the symbolic equations for the linear four-terminal network in y parameters. Its diagram is given in Figure 1. Examining this equivalent diagram, we may predict the properties of the solution of Equations (22) and (23).

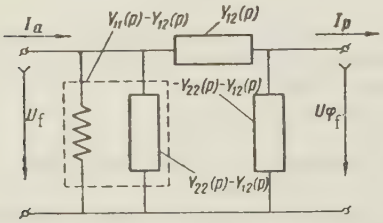


Figure 1. Equivalent diagram of self-excited oscillator for fluctuations

First, it is immediately evident that fluctuation in amplitude and phase are determined by both components of noise current I_a and I_p provided identity $Y_{12}(p) \equiv 0$ does not hold. If this condition is fulfilled. If this condition is fulfilled, Equations (22) and (23) become independent:

$$Y_{11}(p)U_f = I_a, \tag{24}$$

$$Y_{22}(p)U\varphi_f = I_p, \tag{25}$$

and the equivalent circuit in Figure 1 is, in effect two circuits. Moreover, both from the circuit and from Equations (24 and (25), it is evident that fluctuations in amplitude will be determined only by current I, while fluctuations in phase will be determined only by I_p.

Next, it is evident that the circuit in Figure 1 is asymmetrical. Its input (at the side to which voltage U_f is applied) is shunted by a conductance g. In order to appreciate the significance of this asymetry let us examine admittances Y22(p) and $Y_{12}(p)$. From (19), (12) and (21), (13) it is evident that with $p \to 0$ $Y_{22}(p) \to 0$ and $Y_{12}(p) \to 0$. This means that in the low-frequency region both admittances are capacitive in nature. Hence from the input side (to which voltage $U_{\varphi f}$ is applied) with current feed there occurs a charge storage and the mean square $U_{\varphi f}$ will increase indefinitely. It is useful to compare this result with that obtained in [7]. It is not new and indicates only that the random deviations in phase caused by the action of noise are not reestablished. By analogy with the simplest case, we may expect a diffused pattern of phase excursion in the self-excited oscillator. Such charge storage cannot exist at the side to which voltage U_f is applied, for there occurs a dissipation of g. Hence the mean square value of fluctuation is also limited.

Finally, the same equivalent diagram (Figure 1) easily permits determining the expression for the power spectra of voltages U_f and $U_{\varphi f}$ in terms of the power spectra of currents I_a and I_p. For this purpose it is necessary to calculate the complex frequency characteristics of the transfer resistances from I_a and I_p to U_f and $U_{\varphi f}$. Let us first find the symbolic expressions of U_f and $U_{\varphi f}$ in terms of currents I_a and Ip. These may be obtained either directly from the circuit in Figure 1 or by solution of equations (22) and (23). Let us represent the determinant of this system as

$$\Delta(p) = Y_{11}(p) Y_{22}(p) + Y_{12}^2(p). \tag{26}$$

We write the expressions for U_f and $U_{\varphi f}$ in the form

$$U_f = \frac{Y_{22}(p) I_a - Y_{12}(p) I_p}{\Delta(p)},$$
$$U_{\varphi f} = \frac{Y_{11}(p) I_p + Y_{12}(p) I_a}{\Delta(p)}. \tag{27}$$

Replacing p in (27) with $i\Omega$, we obtain the required complex frequency characteristics.

Let us note that if in the first of the equations in (27) we assume $I_a = 0$ and $I_p = 0$ we may derive the differential equation ($p = d/dt$) relative to U_f. In it U_f must be regarded as a small dusturbance in amplitude relative to the stationary value. If we further discard U_f and consider p as the characteristic index, we then obtain the characteristic equation for the investigation of the local stability of the stationary operating condition

$$\Delta(p) = 0. \tag{28}$$

This expression is another way of writing the general characteristic equation for the case of lagless self-bias obtained by Hsieh Hsi in [8].

If we are dealing with shot noise, components I_a and I_p are uncorrelated and their mutual spectrum is equal to zero. Hence, the power spectra of $\overline{U_{f\Omega}^2}$ and $\overline{\varphi_{f\Omega}^2}$ are obtained by superimposing the power spectra of $\overline{I_{a\Omega}^2}$ and $\overline{I_{p\Omega}^2}$ transformed by the frequency characteristics of the corresponding transfer resistances:

$$\overline{U_{f\Omega}^2} = \frac{|Y_{22}(i\Omega)|^2 \overline{I_{a\Omega}^2} + |Y_{12}(i\Omega)|^2 \overline{I_{p\Omega}^2}}{|\Delta(i\Omega)|^2}, \tag{29}$$

$$\overline{\varphi_{f\Omega}^2} = \frac{1}{U^2} \frac{|Y_{11}(i\Omega)|^2 \overline{I_{p\Omega}^2} + |Y_{12}(i\Omega)|^2 \overline{I_{a\Omega}^2}}{|\Delta(i\Omega)|^2}. \tag{30}$$

Moreover, as shown in [3], for shot noise it may be assumed that $\overline{I_{a\Omega}^2} = \overline{I_{p\Omega}^2} = \overline{I_{n\Omega}^2}$ where $\overline{I_n^2\,\Omega}$ is the spectral density of shot noise, and we may further simplify formulas (29) and (30).

Knowing the fluctuation spectra of (29) and (30), we may calculate the mean-square fluctuations of amplitude and frequency:

$$\overline{U_f^2} = \int_0^\infty \overline{U_{f\Omega}^2}\, d\Omega, \tag{31}$$

$$\overline{\omega_f^2} = \overline{\left[\frac{d\varphi_f}{dt}\right]^2} = \int_0^\infty \Omega^2 \overline{\varphi_{f\Omega}^2}\, d\Omega. \tag{31}$$

The formulas for the spectra also permit calculation of the mean-square fluctuation excursion in phase

$$\overline{\Delta\varphi_\tau^2} = \overline{[\varphi(t) - \varphi(t-\tau)]^2}.$$

The expression for $\overline{\Delta\varphi_\tau^2}$ in terms of $\overline{\varphi_{f\Omega}^2}$ is obtained by the method proposed in the report by G.A. Yelkin and M.I. Rodak [9]. The difference spectrum $[\varphi(t) - \varphi(t-\tau)]$ is obtained at the output of the linear four-terminal network consisting of a delay line and a subtracter (Figure 2). Its transfer function has the form

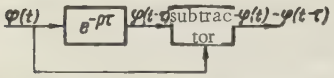


Figure 2. For calculation of the mean-square phase excursion of self-oscillation during time τ.

$$K(p) = 1 - e^{-p\tau}.$$

From this we easily obtain the square of the modulus of the frequency characteristic of this four-terminal network

$$|K(i\Omega)|^2 = 2(1 - \cos\Omega\tau).$$

Thus, for the mean-square of the fluctuation excursion of phase we obtain

$$\overline{\Delta\varphi_\tau^2} = \int_0^\infty \overline{\varphi_{f\Omega}^2} |K(i\Omega)|^2\, d\Omega = 2\int_0^\infty \overline{\varphi_{f\Omega}^2}(1 - \cos\Omega\tau)\, d\Omega. \tag{32}$$

For any self-excited oscillator having an oscillatory system with lumped constants the phase fluctuation spectrum (30) may be written in the form

$$\overline{\varphi_{f\Omega}^2} = \frac{H}{\Omega^2} + \frac{A(\Omega^2)}{B(\Omega^2)}, \tag{33}$$

where H is a certain constant; $A(\Omega^2)$ and $B(\Omega^2)$ are polynomials of Ω^2, wherein the degree of $A(\Omega^2)$ is lower than that of $B(\Omega^2)$ and $B(0) \neq 0$. If we insert (33) into (32) and integrate for Ω we obtain

$$\overline{\Delta\varphi_\tau^2} = \pi H\tau + \psi(\tau), \tag{34}$$

where $\psi(\tau)$ is a certain function exponentially approaching zero with increasing τ. From (34) it is evident that the phase diffusion coefficient D is easily expressed in terms of constant H:

$$D = \pi H. \tag{35}$$

From expressions (33) and (35) we obtain an extremely simple formula for calculation of the phase diffusion coefficient of self-oscillation according to the formula for the phase fluctuation spectrum:

$$D = \pi \lim_{\Omega\to 0} \Omega^2 \overline{\varphi_{f\Omega}^2}. \tag{36}$$

Inserting (30) into (36), we may easily obtain an even simpler equality

$$D = \pi \frac{\overline{I_{p\Omega}^2}}{U^2} \lim_{\Omega\to 0} \frac{\Omega^2}{|Y_{22}(i\Omega)|^2}. \tag{37}$$

Example 1. Double-Tuned Self-Excited Oscillator. Let us calculate the fluctuations in the self-excited oscillator represented in Figure 3. The tuned circuits are considered identical and the coupling loose. Let us introduce the designations:

$$\omega_0 = 1/\sqrt{L(C + C_{cpl}}),$$
$$\beta = \omega_0 C_{cpl} R, \quad y_0 = 1/\beta R,$$
$$T = 2R(C + C_{cpl}),$$
$$p_1 = pT, \quad \nu = \Omega T.$$

Introduction of dimensionless operator p_1 and dimensionless frequency ν decreases the bulk of the calculations. The shortened control admittance, described relative to the reference frequency ω_0, has the form

$$g(p_1) = iy_0[(1 + p_1)^2 + \beta^2]. \qquad (38)$$

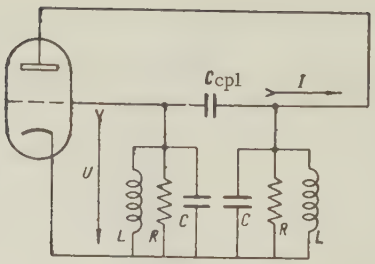


Figure 3. Diagram of double-tuned self-excited oscillator.

Fluctuation calculation begins with definition of the stationary operating conditions. Let us assume that $p_1 = i\alpha$, where $\alpha = \Delta\omega T$, and applying the conditions of (12) and (13) for the steady-state operation, we obtain

$$\alpha = \pm\sqrt{1 + \beta^2}, \qquad (39)$$
$$-2y_0\alpha = S. \qquad (40)$$

It follows from (40) that it is necessary to select the lower sign in (39); that is, the lower of the two natural frequencies of the oscillatory system is selected.

From formulas (19), (20) and (21) we find the elements of the equivalent diagram (Figure 1).

$$Y_{22}(p_1) = -2y_0\alpha p_1,$$
$$Y_{11}(p_1) = -2y_0\alpha p_1 + g, \qquad (41)$$
$$Y_{12}(p_1) = y_0 p_1(p_1 + 2).$$

Let us determine the values of local transconductance σ and coupling factor β at which the self-oscillations are stable, for only at these values of σ and β will the fluctuation calculations have any significance. Substituting $S - \sigma$ for g, in accordance with (18) and (41) we may write the characteristic equation

$$p_1^3 + 4p_1^2 + 4(1 + \alpha^2)p_1 + 2\alpha\left(2\alpha + \frac{\sigma}{y_0}\right) = 0. \qquad (42)$$

Using the Hurwitz criterion, let us determine the region of stability of self-oscillations in plane (β ;σ/y_0) of Figure 4. In this figure the regions of stability for the self-excited oscillator of Figure 3 is delimited by the solid lines. The shaded areas are directed toward the region of stability. The boundary represented by the dashed line in Figure 4 is discussed in Example 2.

Investigation of migration of the roots of the characteristic equation for intersections of the boundaries of the region of stability at the plane (β ;σ/y_0) permits a few observations on the change in the operating conditions of the self-excited oscillator. If the upper boundary is intersected (the real root passes though zero), then there occurs a jump from the discussed operating condition to any other (particularly to the quiescent state). If the lower boundary is intersected (the real part of the complex-conjugate roots passes through zero) then there arises a biharmonic operation.

On the basis of (29), (30) and (41) let us write the expressions for the fluctuation spectra resulting from the shot-noise current. As stated above, we may consider $I^2_{a\nu} = I^2_{p\nu} = I^2_{n\nu}$,

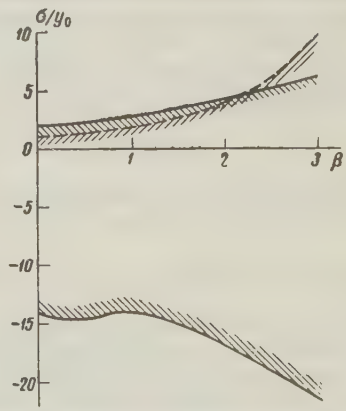


Figure 4. Boundaries of stability: solid line, for self-oscillating circuit of Figure 3; dashed line, same but with phase shift of $-\pi/2$ in feedback circuit.

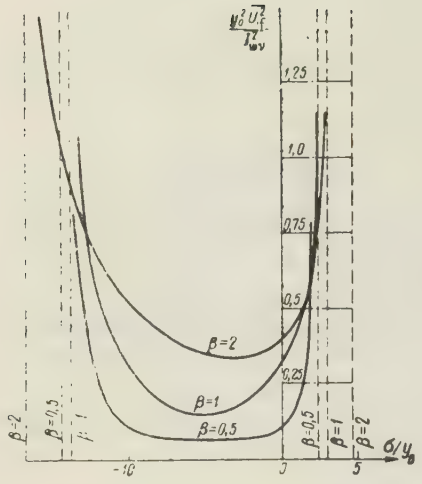


Figure 5. Dependence of mean-square of amplitude fluctuations of double-tuned oscillator of Figure 3 on local transconductance σ/y_0 with different values of coupling β between tuned circuits.

where $I^2_{a\nu}$, $I^2_{p\nu}$ and $I^2_{n\nu}$ are the spectral densities for the dimensionless frequency.

$$\overline{U^2_{f\nu}} = \frac{\overline{I^2_{n\nu}}}{y_0^2}\frac{\nu^2 + 4(1+\alpha^2)}{|\Delta(i\nu)|^2},$$

$$\overline{\varphi^2_{f\nu}} = \frac{\overline{I^2_{n\nu}}}{y_0^2 U^2 \nu^2}\frac{\nu^4 + 4(1+\alpha^2)\nu^2 - \left(2\alpha + \frac{\sigma}{y_0}\right)}{|\Delta(i\nu)|^2},$$

$$|\Delta(i\nu)|^2 = \nu^6 - 8\beta^2\nu^4 + 16\left(1 - \alpha\frac{\sigma}{y_0} + \alpha^4\right)\nu^2 + 4\alpha^2\left(2\alpha + \frac{\sigma}{y_0}\right)^2.$$

From the first of these formulas we may calculate the manner in which the mean-square of amplitude fluctuations within the region of stability varies with a change in local transconductance σ of the oscillatory characteristic and the coupling factor β. These functions are plotted in Figure 5. The ratio of local transconductance of the oscillatory characteristic to admittance y_0 is plotted on the abscissa axis; on the ordinate axis we give the ratio of mean-square amplitude fluctuation to the mean square voltage induced by the noise current for resistances $1/y_0$ in the bandwidth $1/T$.

From these curves it is seen that a sharp increase in fluctuations occurs only in the immediate vicinity of the boundaries of stability. Within this region the mean-square $\overline{U_f^2}$ varies somewhat slowly. With an increase in β the mean-square amplitude fluctuation increases and the change therein within the region of stability becomes more pronounced. In addition, movement along the σ/y_0 axis in the direction of the boundary of stability beyond which the biharmonic regime begins is accompanied by a sharper rise in fluctuations than movement in the direction of extinction of oscillation.

Similar curves may be plotted for the mean-square frequency fluctuation. They will differ from the curves for amplitude in that with movement toward the boundary of oscillation extinction the mean-square frequency fluctuation will tend toward a certain finite value, while with movement toward the boundary of the biharmonic regime the mean-square frequency fluctuation will tend toward infinity.

However, as a rule, these values are of less interest to us than the phase diffusion coefficient, which determines the natural bandwidth of the oscillator circuit. It may be calculated from the formula (37), but it must be kept in mind that we have introduced dimensionless frequency and dimensionless time and that the resulting diffusion coefficient is dimensionless:

$$DT = \pi\frac{\overline{I^2_{n\nu}}}{U^2}\frac{1}{4y_0^2\alpha^2}.$$

Substituting I/S for U in accordance with (16), taking (40) into account, and changing to dimensional time and frequency, we obtain

$$D = \pi\frac{\overline{I^2_{n\Omega}}}{I^2}(\Delta\Omega)^2, \qquad (43)$$

where I is the stationary amplitude of the first harmonic of plate current and $\Delta\Omega = 1/T$ is the bandwidth of one of the tuned circuits.

Expression (43) coincides with the formula for the diffusion coefficient of a single-tuned self-excited oscillator as given in reference [7]. Hence it is necessary that in the given case addition of the second tuned circuit does not change the phase diffusion coefficient. At the end of Example 2 we attempt to explain this result physically.

Example 2. Double-Tuned Self-Excited Oscillator with Phase Inverter. It is of interest to examine the problem of the influence of phase shifts in the feedback circuit on local stability

and fluctuations in the above-discussed self-excited oscillator. Here we consider one special case illustrating the possibility of a considerable change in properties of a double-tuned oscillator by the introduction of phase shift.

Let us assume that in the grid circuit of the self-excited oscillator (Figure 3) there is connected a phase inverter providing a phase shift of $-\pi/2$ and having unit amplitude transfer constant.

Then, instead of (38), with the same designations we obtain

$$y(p_1) = y_0[(1 + p_1)^2 + \beta^2].$$

The conditions of the stationary regime of (12) and (13) yield

$$\alpha = 0,$$
$$S = y_0(1 + \beta^2). \tag{44}$$

The correction for frequency in this self-excited oscillator is equal to zero and the frequency of self-oscillation is equal to the natural frequency of one of the tuned circuits.

The coefficients with all degrees of p_1 are real and, hence,

$$Y_{12}(p_1) = 0, \tag{45}$$

that is, the equivalent circuit in Figure 1 breaks down into two distinct circuits. By means of (19) and (20) let us determine their elements:

$$Y_{11} = y_0 p_1 (2 + p_1) + g,$$
$$Y_{22} = y_0 p_1 (2 + p_1). \tag{46}$$

In accordance with (28), (26) and (44) the characteristic equation is

$$p_1^2 + 2p_1 + \left[1 + \beta^2 - \frac{\sigma}{y_0}\right] = 0. \tag{47}$$

From (47) it is seen that there is only one condition of stability corresponding to the dashed boundary in Figure 4. In transition through this boundary there occurs a jump in amplitude. In the region of large negative values of σ there is, as we see, no second boundary of stability, indicating that the biharmonic regime is not possible in such an oscillator. Detailed analysis of the causes of this is beyond the scope of this report.

Let us proceed to calculation of the fluctuation spectra. In accordance with (29), (30) and (45), (46), we obtain

$$\overline{U_{f\nu}^2} = \frac{\overline{I_{n\nu}^2}}{y_0^2} \frac{1}{\nu^4 + \left(3 - \beta^2 + \frac{\sigma}{y_0}\right)\nu^2 + \left(1 + \beta^2 - \frac{\sigma}{y_0}\right)^2},$$

$$\overline{\varphi_{f\nu}^2} = \frac{\overline{I_{n\nu}^2}}{y_0^2 U^2 \nu^2} \frac{1}{\nu^2 + 4}.$$

Plotting the dependence of the mean-square amplitude and frequency fluctuations on σ/y_0 in this case is not difficult and we shall omit it. Let us calculate only the diffusion coefficient. In accordance with (37), the dimensionless diffusion coefficient is

$$DT = \pi \frac{\overline{I_{n\nu}^2}}{U^2} \frac{1}{4y_0^2}.$$

Replacing U with I/S, taking into account equality (44) and changing to dimensional time and frequency, we obtain

$$D = \pi \frac{\overline{I_{n\Omega}^2}}{I^2} (\Delta\Omega)^2 \frac{1 + \beta^2}{4}. \tag{48}$$

Let us compare the values of the diffusion coefficients in the self-excited oscillators of examples 1 and 2 with identical bandwidths $\Delta\Omega$ of the partial circuits and identical amplitudes of self-oscillation (i.e., identical S).

From the condition of equality of the mean transconductances of (40) and (44), for each β in the circuit of example 1 let us insert two corresponding values of β in the circuit of example 2. One of them is always less than unity. Hence, as follows from (43) and (48), in the circuit of example 2 there may always be obtained a decrease by a factor of 2 - 4 in diffusion coefficient in comparison with the circuit of example 1. This is apparently due to the fact that with $\beta < 1$ the noise bandwidth of the four-terminal feedback network is narrower than the noise bandwidth of the single-tuned circuit. However, in the preceding oscillator with $\beta < 1$ the noise bandwidth is the same but no decrease in diffusion coefficient is observed with a decrease in β. The cause of this lies in the fact that in the oscillator of Figure 3, with a change in β, the frequency of self-oscillation and phase of the feedback coefficient change and the increase in diffusion coefficient caused by their change compensates for the decrease in it due to the decrease in noise bandwidth. In example 2, with a change in β the frequency of self-oscillation does not change, meaning that the phase relationships in stationary operation do not change and hence the gain in diffusion coefficient due to the decrease in noise bandwith appears in pure form.

REFERENCES

1. I.L. Bershteyn, Fluctuations in a self-oscillating system and determination of the natural attenuation of frequency of a vacuum-tube oscillator. ZhTF, 1941, 11, 4, 305
2. I.S. Gonorovskiy, Phase fluctuations in a vacuum-tube oscillator. Dokl. AN SSSR, 1955, 101, 4, 657.
3. S.M. Rytov, Fluctuations in self-oscillating systems of the Thompson type I, II, ZhETF 1955, 29, 3, 304, 315.
4. V.S. Troitskiy, Problems of the theory of fluctuations in self-excited oscillators, Izv. vuzov, MVO SSSR (Radiofizika), 1958, 1, 1, 20
5. A.N. Malakhov, Fluctuations in self-oscillating systems with many degrees of freedom, Izv. vuzov MVO SSSR (Radiofizika), 1958, 1, 2, 71.
6. S.I. Yevtyanov, Relationship between symbolic and shortened equations, Radiotekhnika, 1946, 1, 1, 68.
7. S.I. Yevtyanov, V.N. Kuleshov, Fluctuations in single-tuned self-oscillators, NDVSR (Radiotekhnika i Elektronika), 1958, 1,4, 93.
8. Hsieh Hsi, A general characteristic equation for investigation of stability in the stationary regime of a self-excited oscillator, Radiotekhnika i Elektronika, 1958, 3, 6, 770.
9. G.A. Yelkin, M.I. Rodak, Theoretical investigation of statistical phenomena in self-oscillating systems with automatic frequency control, Report IRE AN SSSR, No 101-02-55.

Submitted to the editors 14 July 1960

ANALOG INVESTIGATION OF SOME OPTIMUM DEVICES FOR FILTERING OF RANDOM-DURATION PULSE SIGNALS

N.K. Kul'man and P.S. Landa

The article reports on investigation of a nonlinear filter for separation of a random-duration pulse signal from white noise on a standard MN-7 analog computer. In order to determine efficiency of the nonlinear filter the results are compared with those obtained

with a linear filter calculated from the Wiener-Kolmogorov theory and simulated on the MN-7 computer. Operation of optimum nonlinear and linear filters in conjunction with threshold computer devices was studied with various signal-to-noise ratios as well as with different signal parameters.

INTRODUCTION

Problems associated with selection of optimum systems achieving separation of weak signals from noise have increasingly greater importance. Theoretical examination of the problem of the optimum method is found in a number of works [1-6]. In the works of Wiener [1] and Kolomogorov [2] there is presented a theory of linear filtering. As is known, if the signal and noise have Gaussian probability distribution, then a linear filter designed in accordance with the Wiener-Kolmogorov theory is the most optimum filter by all criteria. If either the signal or noise does not have Gaussian probability distribution, then the optimum filter system will be a nonlinear filter giving at the output an a posteriori probability distribution for the given type of signal in conjunction with a computing device. (references [3-6]). Choice of the computing device is determined by the assigned optimality criterion of the filter, depending on the specific conditions of the problem.

In references [3,4,], for a posteriori probabilities, integral expressions are presented on the basis of which in many even relatively simple cases it is extremely difficult to construct an actual system. In this respect greater convenience is found in description of the optimum system in the form of differential equations [5], the use of which permits relatively simple implementation of the actual optimum system (for example, by means of electronic analogs.).

In the present work a standard MN-7 analog computer was used for investigation of a nonlinear filter for separation of a pulse signal of random duration from white noise. In order to determine the operational efficiency of the nonlinear filter the obtained results were compared with the results of filtering with a linear system designed according to the Wiener-Kolmogorov theory and simulated on the same MN-7 computer. In a number of special cases theoretical evaluations were obtained for certain quantities characterizing the efficiency of nonlinear and linear filtering which were compared with experimental results. The chosen quantities were: the average number of false pulses per second and the average relative number of undetected pulses as a function of duration. Operation of optimum nonlinear and linear filters was investigated with different signal-to-noise ratios and also with different signal parameters.

1. EQUATION FOR OPTIMUM NONLINEAR AND LINEAR FILTERS.

MODELS OF NONLINEAR AND LINEAR FILTER SYSTEMS.

The signal is a sequence of positive and negative pulses of identical amplitude a, but of random duration (the so-called generalized telegraph signal) with a priori known statistical characteristics: α, the average number of changes from state +a to state -a per second; β, the average number of changes from state -a to +a per second. Such a signal is a unidimensional Markov process. The a priori probabilities $w_1(t)$ and $w_2(t)$ that a signal s at moment t is +a or -a are accordingly subject to the following system of equations:

$$\dot{w}_1 = -\alpha w_1 + \beta w_2,$$

$$\dot{w}_2 = \alpha w_1 - \beta w_2.$$

It is easily shown that quantities α and β are inversely proportional to the mean duration of the positive and negative pulses, respectively.

From relationship (13) in [5] we may obtain the following equation for optimum filtering:

$$\dot{x} = -(\alpha - \beta)\, a - (\alpha + \beta)\, x + \frac{1}{N}(a^2 - x^2)\, r(t), \tag{1}$$

where x(t) is the a posteriori mean value of signal: $x(t) = a(w_{1_{ps}} - w_{2_{ps}})$; $w_{1_{ps}}$ and $w_{2_{ps}}$ are the a posteriori probabilities that the signal has a value of +a or -a, respectively;

$r(t) = s(t) + n(t)$ is the time series "signal + noise" at the filter input; N is the noise intensity, defined by the relationship $\overline{nn_\tau} = N\delta(\tau)$.

For the generation of the sequence of rectangular pulses $\tilde{s}_1(t)$ let us use the criteria of the maximum of a posteriori probability; that is, let us think of a device providing at the output a positive voltage +a if $x(t) > 0$ and a negative voltage -a if $x(t) < 0$. Such a device may be an electronic relay realized on the MN-7 analog computer by one of the schemes of "standard nonlinearities". The filter described by Equation (1), in combination with a threshold device, is an optimum nonlinear filtering system for the given type of signal.

Since the present report is a comparison of the quality of nonlinear and linear filters by use of electronic analogs, it is also convenient to describe the latter by means of a differential equation. For this purpose let us calculate the frequency characteristic of the linear filter in the same manner as in reference [7] for the special case in which the mean duration of positive and negative pulses is equal. Calculation yields

$$\Phi(j\omega) = \frac{8\alpha\beta a^2}{N(\alpha+\beta)^2} \frac{1}{1+\sqrt{1+\frac{8\alpha\beta a^2}{N(\alpha+\beta)^3}}} \frac{1}{(\alpha+\beta)\sqrt{1+\frac{8\alpha\beta a^2}{N(\alpha+\beta)^3}}+j\omega}.$$

It is known that a filter possessing a frequency characteristic of the type is described by the differential equation of $\dot{x} + Ax = Br(t)$, where $r(t)$ is the signal at the filter input and $x(t)$ is the filter output. Inserting the values of A and B, we obtain the following equation for the optimum linear filter:

$$\dot{x} + (\alpha+\beta)\sqrt{1+\frac{8\alpha\beta a^2}{N(\alpha+\beta)^3}}\, x = \frac{8\alpha\beta a^2}{N(\alpha+\beta)^2} \frac{r(t)}{1+\sqrt{1+\frac{8\alpha\beta a^2}{N(\alpha+\beta)^3}}}. \qquad (2)$$

The signal obtained at the output of the linear filter is converted to a series of rectangular pulses $\tilde{s}_2(t)$ also by means of the electronic relay.*

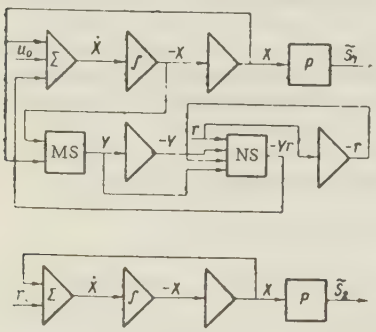


Figure 1. Block diagram of nonlinear and linear filter models: triangle, amplifier-inverter; ∫, integrator: NS, nonlinearity section; MS, multiplier section; Σ, adder; P, electronic relay; $u_0 = (\alpha-\beta)aM_xRC$, where $M_x = X/x$ is the "scale" factor and RC is the integrator time constant.

Figure 1 shows the block diagrams of nonlinear filters with the electronic relay. Behavior of the models is described by the equations for optimum filtering (1) and (2). Voltages X and Y(X) in the models (machine variables) correspond to the quantities x and $y(x) = a^2 - x^2$ in the equations (see [8]).

2. RANDOM-DURATION PULSE GENERATOR AND "WHITE NOISE" GENERATOR. MEASUREMENT OF SPECTRAL DENSITY OF NOISE.

A random-duration pulse generator (Figure 2) serves for the generation of the effective signal s(t) and consists of a low-frequency noise generator and a trigger circuit. The trigger circuit uses two relays of type RSM and one polarized relay of type RP-5.

To the winding of an RP-5 relay (re 3) there is applied a low-frequency random voltage $n_1(t)$ from a noise generator (NG 1) which causes a corresponding displacement of the armature. The result is appearance at the circuit output of a sign which randomly assumes values of +a and -a with probabilities α and β of the order of unity (mean pulse duration of the order of 1 sec). NG 1 is an infrasonic noise generator with the bandwidth of the order of 15 cps (described in [9]).

*The filter system, consisting of a linear filter and relay, will hereafter often be referred to as "linear" despite the fact that it does contain a relay, which is in principle a nonlinear element.

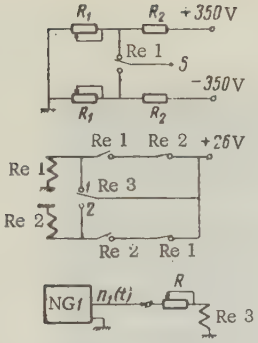


Figure 2. Basic diagram of random-duration pulse generator.

By changing the value of noise applied to Re 3 and displacing its mean value we may change the mean number of alternations in sign of the signal per second (α and β)

In order to obtain wideband low-frequency noise mixed with the signal, a low-frequency noise generator (NG2) similar to NG 1 was made. NG 2 provides at the output a sideband random process $n_2(t)$ which may be mixed with the signal as "white noise" (noise bandwidth of the order of 100 cps).

An adder consisting of an ordinary resistance-coupled DC amplifier serves to generate the time series of signal pulse noise. The effective-signal transfer constant is chosen equal to unity, whereas the spectral density of the noise may be varied over somewhat wide limits by changing the noise transfer constant K. As a result, at the adder output we have a signal $r(t) = s(t) + n(t)$, where $n(t) = Kn_2(t)$. Since the signal s(t) and noise n(t) are usually statistically independent, as was assumed in deriving the equations for optimum filtering, the noise source $n_1(t)$ and $n_2(t)$ must be uncorrelated.

An optimum filter may be constructed only in the case where the properties of the signal and the noise are known a priori. Hence, it is necessary to determine the spectral density of noise N in the operating frequency range. For this purpose it suffices to measure the noise dispersion σ^2 and determine its spectral composition $F(j\omega)$. Then $N = \sigma^2/\Delta f$, where

$$\Delta f = \int_{-\infty}^{\infty} F(j\omega)\, F^*(j\omega)\, d\omega/2\pi\, F^2(0) -$$

is the equivalent bandwidth. The noise dispersion σ^2 is approximately expressed by the formula

$$\sigma^2 = \frac{1}{T}\int_0^T n^2(t)\, dt,$$

if the averaging time T is much greater than the noise correlation time.

The noise dispersion value n(t) was determined experimentally. For this purpose noise was first applied to a squarer and then to an integrator, the output voltage of which was measured. The values of noise intensity and (α and β) for which experimental data were obtained are listed in Table 1

Table 1

No.	Experiment	Q	Mean No. of false pulses per sec.				Mean rel. lag in pulse processing, %						Total rel error, %		Rel. error in processing of duration, %	
			nonlinear system		linear system		nonlinear system			linear system			nonlinear		linear	
			exper. results	theor. est.	exper. results	theor. est.	pos. pulse	neg. pulse		pos. pulse	neg. pulse		syst.		syst.	
1	$N = 0{,}26$ v^{2}sec. $\alpha = \beta = 1$ sec^{-1}	0,13	—	—	—	—	9						11	11	2	4
2	$N = 0{,}26$ " $\alpha = \beta = 2$ "	0,26	—	—	—	—	13						19	20	6	7
3	$N = 0{,}13$ " $\alpha = \beta = 1$ "	0,06	—	—	—	—	7						9	9	2	2
4	$N = 0{,}26$ " $\alpha = 6{,}1$ " $\beta = 0{,}6$ "	0,44	0,4	0,4	2,3	1,7	28	2		21	5		9	15	—	—
5	$N = 0{,}13$ " $\alpha = 6{,}1$ " $\beta = 0{,}6$ "	0,22	0,4	0,4	1,4	1,2	27	2		23	4		7	9	—	—

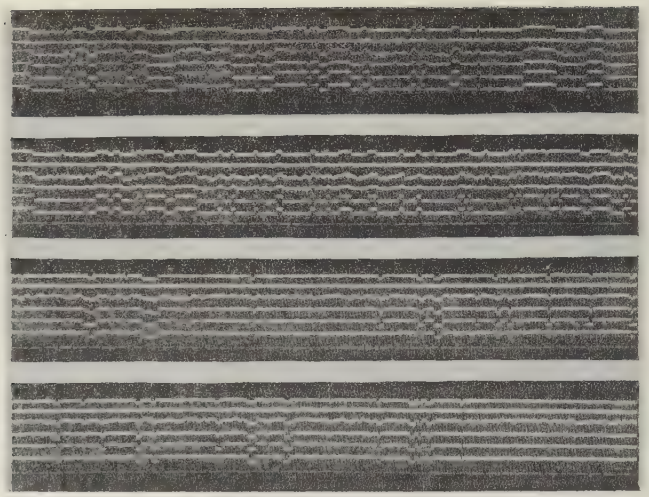


Figure 3. Typical oscillograms. Top to bottom: effective signal s(t); time series r(t)=s(t)+n(t); output voltage of linear filter; output voltage of nonlinear filter; pulse signal $\tilde{s}_2(t)$ from output of linear system; pulse signal $\tilde{s}_1(t)$ from out put of nonlinear system and second markers -- pulses developed by a special section in the MN-7 analog computer (a, experiment 3, b, experiment 2; c, experiment 4, d, experiment 5).

3. RESULTS OF EXPERIMENT

An MPO-2 loop oscillograph was used in recording the results of nonlinear and linear filtering. The oscillograms of each experiment, containing more than a hundred positive and more than a hundred negative effective signal pulses, were quantitavely processed and all the errors arising in filtering of the affective signal were noted. Typical oscillograms are shown in Figure 3.

It must be pointed out that, first of all, in each experiment the effective signal s(t) was studied: on the basis of more than 100 positive and more than 100 negative pulses the average duration and probability of signal change α and β were determined and compared with the a priori given values. The error in determining α and β was approximately 10 - 15 percent. In view of the noncoincidence of the actual constants for the transition in effective signal with a priori assigned values as well as the error in determining the spectral density of noise (of the order of 25 percent), the investigated filter models are not quite optimum, but are sufficiently close to such systems.

For evaluation of efficiency of filter operation various criteria may be introduced, depending on the specific practical requirements for operation of the device in which it is used.

In many cases it is convenient to evaluate filter efficiency from the point of view of total relative time of incorrect processing of the signal-- total relative error. Of course, the practical value of this criterion depends on the specific characteristics of the effective signal. If the effective signal is a phototelegraph signal, then in the simplest case this is a symmetrical signal assuming two values, one of which corresponds to "white" and the other to "black." It is evident that in this case the information is contained in the duration of the pulses and that incorrect processing of it leads to distortion of the picture. The regular delay in pulse detection associated with the lag of the filter system does not cause loss of information, for the mean lag error may be taken into account. Consequently, it is of greater practical interest to evaluate the efficiency of filtering of a symmetrical pulse signal from the point of view of relative error in processing of pulse

duration, which may be defined as the difference of total relative error and the mean relative error in delay.

The results of experimental investigation of filtering of a symmetrical signal are given in Table 1. (Experiments 1-3). As is seen from Table 1, the filtering errors increase with noise intensity. A twofold increase in the mean number of signal transitions per second with the noise constant also leads to a significant increase in filtering error, which gives the impression of intense noise effect on the system. This, too, is understandable, for with an increase in the average number of signal transitions per unit time the signal spectrum is increased and there must be a corresponding increase in the passband of the receiving unit, which causes an increase in the disturbing effect of noise on the results of filtering. Thus, the total relative error increases with an increase in $Q = N(\alpha + \beta)/a^2$, which may be referred to as the generalized signal-to-noise ratio.

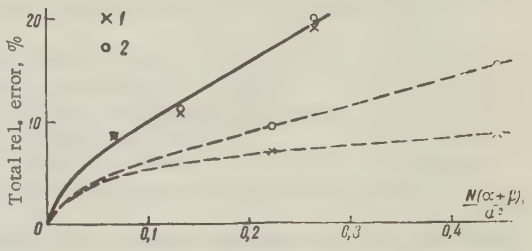


Figure 4. Total relative error in filtering (experiments 1-5): solid line, filtering of a symmetrical signal: dashed lines, filtering of an asymmetrical signal; 1, experimental values for nonlinear filtering; 2, experimental values for linear filtering.

Figure 4 (solid curve) shows the dependence of total relative error on Q for the cases of nonlinear and linear filtering. The total relative error in the case of a symmetrical signal is approximately the same for nonlinear and linear filtering (difference ~1%).

Also of some interest is the case of filtering of an asymmetrical signal. With $\alpha \gg \beta$ (sharply asymetrical signal) the mean duration of the positive pulse is much less than the mean duration of the negative pulse, with the result that the negative pulse may be regarded as a pause of random duration between the positive pulses, which carry the information, and the problem of filtering acquires the character of detection of a random pulse signal in noise.

In this practical case of detection of a positive pulse in the presence of noise we are dealing with two types of error: errors of nondetection, wherein the system does not react to a change in polarity of signal; errors of false detection, wherein the system indicates a change in signal polarity which has not occurred. In view of the fact that a priori knowledge of these errors is often required before designing filter systems, it is of particular interest to attempt to evaluate them, even if only within a certain approximation. For the case of filtering of a sharply asymetrical signal ($\alpha \gg \beta$) we may obtain theoretical estimates for the average number of spurious signals per unit time and the probability of nondetection of the positive pulse depending on its duration. Theoretical estimates were obtained in the case of $\alpha = 6.1$ sec^{-1} and $\beta = 0.6$ sec^{-1} for different values of noise intensity. The principal results of experiments for an asymmetrical signal with the derived theoretical estimates are also listed in Table 1 (Experiments 4 and 5). As is seen from Table 1, with an increase in Q there is an increase in error of false detection and an increase in lag error.

For an asymmetrical signal at any noise level, the average number of false pulses per unit time is much smaller in nonlinear filtering than in linear filtering. It is true that there is also a greater lag in detection of positive pulses (Figure 5), but this error is a relatively small part of the total filtering error, as a result of which positive pulses are infrequent. The greatest role is played by error due to the occurrence of false pulses in the long intervals between positive signal pulses. Figure 4 (dashed curves) shows the dependence of total relative error on Q. As is seen from Figure 4, the total relative error for a nonlinear system is considerably smaller than for a linear system.

For the purpose of improving the operation of a linear system, at the output of the linear filter in experiment 6, a relay with variable operating threshold b was used. Principal results of the experiment are given in Table 2. The values of α, β, and N are the same as in experiment 4. As is seen from Table 2, with an increase in the operating threshold there is a decrease in the average number of false pulses per second but the nondetection error increases. At the optimum threshold (b = 0.4 to 0.6 v) operation of the filter system is almost the same as that of the corresponding nonlinear system (investigated in experiment 4) with regard to the total relative filtering error.

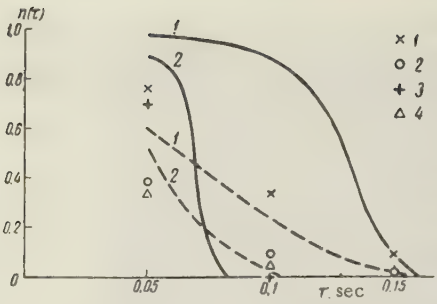


Figure 5. Probability of non detection of positive pulses as a function of their duration (experiments 4 and 5): theoretical estimate given in the form of a curve (1 experiment 4; 2, experiment 5); solid curve , nonlinear filtering; dashed curve, linear filtering; 1, experimental value of probability of nondetection for the nonlinear system; 2, the same for the linear system in experiment 4; 3 experimental value of probability of nondetection for the nonlinear system; 4, the same for the linear system in experiment 5.

However, the following circumstance is of extreme importance: at the optimum operating threshold short positive pulses are detected by the modified linear system with greater relative lag than in the case of nonlinear filtering, whereas the number of false pulses is approximately the same. If we consider that the long pulses correspond to the interval between messages and that the short pulses are the information carriers (for example, the control signals in remote control systems), then the nonlinear system of filtering has the advantage over the modified linear system.

Table 2

b, v	Exper. value of av. no. of false pulses per sec.	Theor. value of av. no. of false pulses per sec.	Av. rel. delay in detection of pulses, %		Total rel. error, %
			pos.	neg.	
0.12	2,0	1,2	20	5	12
0.2	1,5	0,9	34	4	11
0.4	0,7	0,4	35	2	8
0.6	0,4	0,1	54	1	8
0.64*	—	—	100	0	9

With b = 0.64 v detection of positive pulses generally ceased and the total relative error of linear filtering was 9 %, a value equal to the relative content of positive pulses in the effective signal.

Note: comma denotes decimal point.

CONCLUSIONS

1. The total relative error in the filtering increases monotonically with an increase in the value of $Q = N(\alpha+\beta)/a^2$ of the generalized signal-to-noise ratio whereby for a fixed value of Q the total relative error both for a nonlinear and for a linear system in the case of filtering of an asymmetrical signal is smaller than the corresponding value of error in filtering a symmetrical signal.

2. In evaluating the results of filtering form the point of view of total relative error we may assert that in the case of separation of an asymmetrical effective signal from noise the nonlinear system is clearly superior. In the case of filtering of a symmetrical signal the nonlinear and linear filters yield approximately the same error.

3. In the presence of an asymmetrical signal at the filter input the operation of a linear system may be improved by selecting the appropriate operating threshold for a relay at the linear filter output, so that the modified linear system will yield approximately the same total relative error as a nonlinear system.

4. Theoretical estimates of the errors of false detection and nondetection as calculated for a number of specific cases gave wholly satisfactory agreement with experimental results.

In conclusion the authors wish to thank S. P. Strelkov and R. L. Stratonovich for their interest in this work and for their valuable advice.

REFERENCES

1. N. Wiener, The Extrapolation, Interpolation and Smoothing of Stationary Time Series, J. Wiley, N.Y., 1949.

2. A.N. Kolmogorov, Interpolation and Extrapolation of Stationary Random Series, Izvestiya, AN SSSR Ser. matem., 1941, 5, 3.
3. W.D. White, The Role of Nonlinear Filters in Electronic Systems, National Electronics Conference Proc., 1954, 9, (Chicago).
4. I.N. Amiantov, Primeneniye teorii resheniy k zadacham obnaruzheniya signalov i vydeleniya signalov iz shumov. [Application of the Theory of Solutions to Problems of Detection of Signals and Separation of Signals from Noise]. Izd. VVIA im. N. Ye. Zhukovskiy, 1958.
5. R.L. Stratonovich, Toward a Theory of Optimum Nonlinear Filtering of Random Functions, Teoriya veroyatnostey i yeye primeneniya, 1959, 4, 2, 239.
6. F.M. Woodward, Probability theory and Information Theory with Radar Applications. translated from English, Izd, Sovetskoye radio, 1955.
7. V.V. Solodovnikov, Vvedeniye v statisticheskuyu dimaniku sistem avtomaticheskogo regulirovaniya [Introduction to the statistical dynamics of automatic control systems], GTI, 1952, p. 273.
8. B. Ya Kogan, Elektronnyye modeliruyushchiye ustroystva i ikh primeneniye dlya issledovaniya sistem avtomaticheskogo regulirovaniya, [Electronic analog computers and their application in investigation of automatic control systems]. Fizmatgiz, 1959.
9. Yu. M. Romanovskiy, Low-frequency noise generators, Pribory i tekhnika eksperimenta, 1958, 4, 98.

Physics Department
M.L. Lomonosov
Moscow State University

Submitted to the editors 22 April 1960

DETECTION OF RANDOM SIGNALS AGAINST A NOISE BACKGROUND

V.F. Pisarenko

The report discusses the problem of detection of a random signal against a noise background with continuous observation during a finite segment of time. For the given class of signals and noise, methods are presented for determining the limiting likelihood ratio and its characteristic function.

INTRODUCTION

We shall examine the problem of detection of a signal against a noise background in the case where the signal (as well as the noise) is a random process. Up to the present time the examination of problems in detection of a signal (whether random or nonrandom) has usually dealt only with the values of the observed processes in a few discrete moments in time (see reference [1]). As a rule, it has been possible to calculate the detection characteristics in clear form only in the case where the values of the observed process at the chosen moments may be considered to be independent.

In distinction from this approach to solution of the presented problem we will use all values of the observed process over a certain segment of time (compare with reference [2]). Generally speaking, for practical purposes it is always possible to limit examination to a finite number of observation points which are sufficiently close together. However, in such discrete observation of correlated quantities there arises a number of difficulties in calculation.

For example. if a process is observed at moments $t_1 \ldots . t_n$, then for the problem under discussion we must determine the eigenvalues of the n-th order matrices and transform these matrices, which, as is known, is no easy problem in calculation. For continuous observation the problem reduces to finding the eigenvalues of a few integral equations and solving these equations which can be done in the case of the classes of signal and noise discussed below. It is evident that in continuous observation the information contained in the representation of the process is more fully used than in discrete observation; however, we shall not discuss this problem, if only for the reason that determining the characteristics of discrete observation of correlated quantities, as has already been stated, is no easy problem in calculation. Let it be noted only that if the distance between points of discrete observation is much smaller than the intervals of signal and noise correlation, it may be considered that the real characteristics of detection for discrete and continuous observation are identical.

Let us formulate in slightly more general form than indicated in the title the problem of testing two hypotheses concerning the random process. A special case of this problem is that of detection of the signal against a noise background.

Let us repeatedly observe over the time interval [O, T] a random process $\xi(t)$, wherein each observation of two cases is possible: either $\xi(t) = \eta_0(t)$ or $\xi(t) = \eta_1(t)$ where $\eta_0(t)$ and $\eta_1(t)$ are stationary Gaussian processes with a mean zero* and rational spectral density. This indicates that the Fourier transform of the correlation functions $\eta_0(t)$ and $\eta_1(t)$, designated by K(t-s) and L(t-s) respectively, has the form

$$\left|\frac{Q_0(i\lambda)}{P_0(i\lambda)}\right|^2 \text{ and } \left|\frac{Q_1(i\lambda)}{P_1(i\lambda)}\right|^2,$$

where $Q_0(i\lambda)$, $Q_1(i\lambda)$, $P_0(i\lambda)$ and $P_1(i\lambda)$ are polynomials without roots in the lower half-plane of complex variable λ (from the engineering standpoint) processes with rational spectral density may be obtained by passage of "white noise" through a circuit with lumped constants).

Knowing the sample function of process $\xi(t)$ over time segment [0, T], we shall test two mutually exclusive hypotheses regarding this sample function: hypothesis H_0 assumes that we have observed the sample function of the process $\eta_0(t)$; hypothesis H_1 assumes that we have observed the sample function of the process $\eta_1(t)$. In testing hypotheses H_0 and H_1 let us use a criterion based on the likelihood ratio (see reference [1]) (more precisely, based on a certain generalization of this concept for the case where the observed process cannot be characterized by a finite number of random parameters (see reference [3]).

The usual scheme of signal detection against a noise background (signal and noise independent) corresponds to the case in which L(t-s) = K(t-s) + M(t-s), where K(t-s) is the correlation function of noise and M(t-s) is the correlation function of signal.

Section 1 of this report discusses a method of determining the limit of the likelihood ratio. Section 2 discusses a method of determining the characteristic function of the limiting likelihood ratio. Section 3 evaluates L. A. Vaynshteyn's remarks on a more realistic statement of the problem of signal detection in the presence of noise. The specific formulas derived in sections 1 and 2 are apparently new; however, their derivation is not quite correct from the mathematical standpoint and, hence, a rigid mathematical basis for these formulas remains to be presented.

1. DETERMINATION OF THE LIMIT OF THE LIKELIHOOD RATIO

In the discussion of continuous random processes it is convenient to change from the random process $\xi(t)$ to the discrete counting system of random values of $\{\zeta_n\}$ $(n = 1, 2 \ldots)$ associated with $\xi(t)$, which values we shall refer to as coordinates of the process. It is also necessary to choose ζ_n so that all the information concerning $\xi(t)$ is retained in $\{\zeta_n\}$. For example, if by $\{\zeta_n\}$ we take the values of $\xi(t)$ at all rational points of segment [O, T], then for continuous $\xi(t)$ the representation of the values of $\{\zeta_n\}$ is equivalent to the representation of $\xi(t)$. It is often convenient to choose as $\{\zeta_n\}$ the Fourier coefficients

*It is shown below that derived results are easily transferred to the case of arbitrary known mean values.

of the process $\xi(t)$ for some system of functions of $\{\varphi_n(t)\}$ which is defined on $[0, T]$:

$$\zeta_n = \int_0^T \xi(t)\,\varphi_n(t)\,dt.$$

For example, an ordinary system of trigonometric functions at $[O, T]$ could be taken as $\{\varphi_n(t)\}$.

Thus, we shall consider that in this manner we have changed from the continuous process $\xi(t)$ to its coordinates $\{\zeta_n\}$. If we designate the joint probability density of the first m coordinates of $\zeta_1 \ldots \zeta_m$ with the hypotheses H_0 and H_1, respectively, as $p_0(y_1 \ldots y_m)$ and $p_1(y_1 \ldots y_m)$, then the usual likelihood for $\zeta_1 \ldots \zeta_2$ is written as:

$$l_m(y_1 \ldots y_m) = p_0(y_1 \ldots y_m) / p_1(y_1 \ldots y_m).$$

Let us insert into $l_m(y_1 \ldots y_m)$ the random vector $(\zeta_1 \ldots \zeta_m)$, which is represented as $(x_1 \ldots x_m)$

$$l_m(x_1 \ldots x_m) = p_0(x_1 \ldots x_m) / p_1(x_1 \ldots x_m). \tag{1}$$

Since with an increase in m the information contained in $\zeta_1 \ldots \zeta_m$, then we naturally convert in ratio (1) to the limit with $m \to \infty$. Although densities $p_0(x_1, \ldots x_m)$ and $p_1(x_1 \ldots x_m)$ with $m \to \infty$ do not, generally speaking, tend toward any limits, their ratio $m(x_1 \ldots x_m)$ with certain additional bounds on K(t-s) and L(t-s) tends toward a definite limit which depends on the sample function x(t). This limit, which we shall designate as F[x(t)] as shown in reference [3], is also that functional of the sample function which generalizes the concept of the usual likelihood ratio for the infinite-dimensional case.*

Note. If the condition

$$\lim_{\lambda \to \infty} \left| \frac{Q_0(i\lambda)\,P_1(i\lambda)}{P_0(i\lambda)\,Q_1(i\lambda)} \right|^2 \neq 1, \tag{2}$$

is fulfilled, then, as is known (see reference [4]), we may determine without error at any time of observation T which of the two competing hypotheses is true (this is discussed in greater detail in section 3). In addition, the functional F[x(t)] assumes the value 0 in all sample functions of the process $\eta_0(t)$ and the value ∞ in all sample functions of the process $\eta_1(t)$. Conversely, upon satisfying the condition

$$\lim_{\lambda \to \infty} \left| \frac{Q_0(i\lambda)\,P_1(i\lambda)}{P_0(i\lambda)\,Q_1(i\lambda)} \right|^2 = 1 \tag{3}$$

it may be shown that functional F[x(t)] everywhere assumes finite values. Let us demonstrate how the functional F[x(t)] may be found in explicit form in case of rational spectral densities and upon fulfilling condition (3).

As $\{\zeta_n\}$ let us take the values of $\xi(t)$ at points of the form $t_{nm} = Tm/n$, $m = 1, 2 \ldots ; n$, $n = 1, 2 \ldots$. Let us express the matrices $\{K(t_{nj} - t_{ni})\}$ and $\{L(t_{nj} - t_{ni})\}$ in terms of K_n and L_n and the corresponding inverse matrices in terms of K_n^{-1} and L_n^{-1}. In the case discussed by us the matrices of K_n and L_n are everywhere nondegenerate, hence the inverse matrices of K_n^{-1} and L_n^{-1} exist.

The probability densities of random values of $\xi(t_{n1}) \ldots \xi(t_{nn})$ when hypotheses H_0 and H_1 are true are written as follows:

---

*In reference [5] it is demonstrated that for each process $\eta(t)$ there may be compared a measure determined over a minimum Borel body B of subsets of the space of real functions, which contains all cylindrical sets. If dimensions μ_0 and μ_1, corresponding to processes $\eta_0(t)$ and $\eta_1(t)$, are such that μ_0 is absolutely continuous relative to μ_1 then the generalization of the likelihood ratio for the infinite-dimensional case will be a Radon-Nikodim derivative of dimension μ_0 in terms of μ_1 (see [3]) which we have designated as F[x(t)]. In the infinite-dimensional case the Radon-Nikodim derivative is the usual likelihood ratio.

$$p_0(y_1 \dots y_n) = (2\pi)^{-\frac{n}{2}} |K_n|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\sum_{i,j} \varkappa_{ij}^n y_i y_j\right\},$$

$$p_1(y_1 \dots y_n) = (2\pi)^{-\frac{n}{2}} |L_n|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\sum_{i,j} \nu_{ij}^n y_i x_j\right\}.$$

Here $\varkappa_{ij}^n$ and ν_{ij}^n are elements of matrices K_n-1 and L_n-1. Let us next obtain the usual likelihood ratio for random values of $\xi(t_{n1}) \dots (\xi\, t_{nn})$ and insert into it the values of sample function $x_{ni} = x(t_{ni})$:

$$l_n(x_{n1} \dots x_{nn}) = |L_n K_n^{-1}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\sum \varkappa_{ij}^n x_{ni} x_{nj} + \frac{1}{2}\sum \nu_{ij}^n x_{ni} x_{nj}\right\}. \tag{4}$$

Let us designate $\sum_j \varkappa_{ij}^n x_{nj} = u_{ni}/n$, $i = 1, 2, \dots n$. Examining these equalities as a system of linear equations relative to x_{ni} and solving it, we obtain

$$x_{nj} = \sum_i K(t_{nj} - t_{ni}) \frac{u_{ni}}{n}, \quad j = 1, 2, \dots, n. \tag{5}$$

We further have

$$\sum_{i,j} \varkappa_{ij}^n x_{ni} x_{nj} = \sum_i \frac{x_{ni} u_{ni}}{n}.$$

Similar operations may be performed for the second term in the exponent of Equation (4). Now let us complete the formal transition to the limit with $n \to \infty$. Then equality (5) the right-hand side of which may be regarded as an integral sum with the interval of 1/n, goes over into the integral equation

$$x(t) = \int_0^T K(t-s)\, u(s)\, ds, \quad 0 \leqslant t \leqslant T.$$

In the same manner the sum $\sum_i \frac{x_{ni} u_{ni}}{n}$ becomes the integral $\int_0^T x(t)\, u(t)\, dt$.

For the second addend in the exponent of (4) we similarly obtain instead of the sum $\sum_{ij} \nu_{ij}^n x_{ni} x_{nj}$ the integral $\int_0^T x(t)\, v(t)\, dt$, where v(t) is the solution of the integral equation

$$x(t) = \int_0^T L(t-s)\, v(s)\, ds, \quad 0 \leqslant t \leqslant T.$$

Thus, as the result we obtain

$$\lim_{n\to\infty} l_n(x_{n1} \dots x_{nn}) = \lim_{n\to\infty} \left|\frac{L_n}{K_n}\right|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\int_0^T x(t)\,[u(t) - v(t)]\, dt\right\}, \tag{6}$$

where u(t) and v(t) are solutions of the integral equations

$$\left.\begin{aligned} &\int_0^T K(t-s)\, u(s)\, ds = x(t), \\ &\int_0^T L(t-s)\, v(s)\, ds = x(t) \end{aligned}\right\} 0 \leqslant t \leqslant T. \tag{7, 8}$$

Equations of the form of (7) and (8) are encountered in the theory of optimal linear prediction of random processes. As is shown in references [6, 12], for the case of rational spectral densities the solution of Equation (7) must be sought in the form

$$u(s) = \tilde{u}(s) + \sum_{k=0}^{l-1} a_k \delta^{(k)}(s) + \sum_{k=0}^{l-1} b_k \delta^{(k)}(s - T),$$

where $\tilde{u}(s)$ is the solution of the differential equation

$$Q_0\left(\frac{d}{dt}\right)Q_0\left(-\frac{d}{dt}\right)\tilde{u}(t) = \frac{1}{2\pi}P_0\left(\frac{d}{dt}\right)P_0\left(-\frac{d}{dt}\right)x(t), \quad 0<t<T; \tag{9}$$

a_k and b_k are constants $\delta^{(k)}(s)$ is the k-th derivative of the delta-function; l is the difference in degree of the polynomials $P_0(i\lambda)$ and $Q_0(i\lambda)$. In order to determine constants a_k, b_k and the constants entering into the general solution of Equation (9), we may simply insert u(s) with all the undefined constants into (7), integrate the expression under the integral sign and equate to zero the coefficients of the linear-independent functions. The undefined constants are found from the resulting equations. Similarly, the function v(s) is found from Equation 8. In the example given below we use a slightly different method of determining the undefined constants.

Despite the fact that the solutions of Equations (7) and (8) contain σ-functions and their derivatives as well as derivatives of x(t), which will be generalized functions [since x(t) has only an $(l-1)$ ordinary derivative, while the expression $P_0(d/dt)P_0(-d/dt)x(t)$ contains derivatives of x(t) of higher order], if we insert these solutions into (6), many of the generalized functions cancel out. If we perform several integrations by parts, the result is that expression (6) may be reduced to a form containing only ordinary (not generalized) functions. It will be noted that in this case condition (3) plays a large part in this procedure.

Let us present in detail a specific example of the described method of determining F[x(t)]. Let $K(t-s) = \frac{1}{\alpha}e^{-\alpha|t-s|}$ and $L(t-s) = \frac{1}{\beta}e^{-\beta|t-s|}$. Equation (7) has the form

$$\frac{1}{\alpha}\int_0^T u(s)\,e^{-\alpha|t-s|}\,ds = x(t).$$

Inserting into Equation (7) u(s) in the form

$$u(s) = \tilde{u}(s) + a\delta(s) + b\delta(T-s).$$

we obtain

$$\frac{1}{\alpha}\int_0^T \tilde{u}(s)\,e^{-\alpha|t-s|}\,ds + \frac{a}{\alpha}e^{-\alpha t} + \frac{b}{\alpha}e^{-\alpha(T-t)} = x(t). \tag{10}$$

Then, instead of $\frac{1}{\alpha}e^{-\alpha|t-s|}$, let us insert $\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{e^{i\lambda(t-s)}d\lambda}{\alpha^2+\lambda^2}$ and apply to both sides of the equality the differential operator $\frac{-d^2}{dt^2} + \alpha^2$. We obtain on the left hand side of the equality

$$\frac{1}{\pi}\int_0^T \tilde{u}(s)\,ds\int_{-\infty}^{\infty}\frac{\alpha^2+\lambda^2}{\alpha^2+\lambda^2}e^{i\lambda(t-s)}\,d\lambda = 2\int_0^T \tilde{u}(s)\,\delta(t-s)\,ds = 2\tilde{u}(t).$$

Consequently, $\tilde{u}(t) = -\frac{x''(t)}{2} + \frac{\alpha^2}{2}x(t)$. In order to determine a and b let us apply to (10) the operator $\frac{d}{dt} + \alpha$ and let $t \to T$. We obtain

$$\frac{1}{\pi}\int_0^T \tilde{u}(s)\,ds\int_{-\infty}^{\infty}\frac{e^{i\lambda(T-s)}}{\alpha - i\lambda}\,d\lambda + 2b = x'(T) + \alpha x(T).$$

Since T-s in the last integral is positive with $0<s<T$, according to the Cauchy theorem this integral is equal to zero. Consequently

$$b = \frac{1}{2}x'(T) + \frac{\alpha}{2}x(T).$$

Applying the operator $-\frac{d}{dt} + \alpha$ to (10) and letting $t \to 0$, we find that

$$a = -\frac{1}{2}x'(0) + \frac{\alpha}{2}x(0).$$

As a result

$$u(s) = \frac{1}{2}\{-x''(t) + \alpha^2 x(t) + \delta(t)[-x'(0) + \alpha x(0)] + \\ + \delta(t-T)[x'(T) + \alpha x(T)]\}.$$

Replacing α with β in this equation we find

$$v(s) = \frac{1}{2}\{-x''(t) + \beta^2 x(t) + \delta(t)[-x'(0) + \beta x(0)] + \\ + \delta(t-T)[x'(T) + \beta x(T)]\}.$$

Inserting the right hand sides of the equations for u(s) and v(s) into formula (6) we obtain

$$F[(x,t)] = \lim_{n\to\infty}\left|\frac{L_n}{K_n}\right|^{\frac{1}{2}} \exp\left\{\frac{\beta^2-\alpha^2}{4}\int_0^T x^2(t)\,dt + \frac{\beta-\alpha}{4}[x^2(0)+x^2(T)]\right\}. \tag{11}$$

The limit before the exponential is of no interest, for it does not depend on the function x(t) (concerning the determination of this limit in explicit form, see Section 2). Expression (11) could have been obtained by another method. Since for the correlation function $K(t-s) = \frac{1}{\alpha}e^{-\alpha|t-s|}$ at any n it is not difficult to write the inverse matrix K_n^{-1}, we may immediately calculate the limit of the quadratic form of the exponential of (4) (see references [7, 13]). The result is that for F[x(t)] expression (11) is again obtained. For the case where $Q_0(i\lambda) \equiv 1$, it was possible to obtain formulas for the undefined constants entering into the solutions of Equation (7). The result is that the solution of Equation (7) has the form

$$u(t) = \frac{1}{2\pi}P_0\left(\frac{d}{dt}\right)P_0\left(-\frac{d}{dt}\right)x(t) + \\ + \frac{1}{2\pi}\sum_{l=0}^{n-1}\delta^{(l)}(t)\sum_{m=0}^{2n-1-l}x^{(m)}(0)\sum_{k=0}^{m}\pi_k\pi_{m+l+1-k}(-1)^k + \\ + \frac{1}{2\pi}\sum_{l=0}^{n-1}\delta^{(l)}(t-T)\sum_{m=0}^{2n-1-l}x^{(m)}(T)\sum_{k=0}^{m}\pi_k\pi_{m+l+1-k}(-1)^{m+l-k},$$

where π_k is the coefficient at z^k in polynomial $P_0(z)$, whence there is immediately derived an expression for F[x(t)] if $Q_0(i\lambda) \equiv 1$ and $Q_1(i\lambda) \equiv 1$.

In this case

$$F[x(t)] = \lim_{n\to\infty}\left|\frac{L_n}{K_n}\right|^{\frac{1}{2}} \exp\left\{-\frac{1}{4\pi}\left[\sum_{k=0}^{n-1}(P_{2k}-\tilde{P}_{2k})(-1)^k \times \right.\right. \\ \left.\left. \times\int_0^T \{x^{(k)}(t)\}^2\,dt + M + N_T - N_0\right]\right\},$$

where

$$M = \sum_{i,j=0}^{n-1}[(-1)^i x^{(i)}(T)x^j(T) + (-1)^j x^{(i)}(0)x^{(j)}(0)] \times \\ \times \sum_{k=0}^{i}(-1)^k[\pi_k\pi_{i+j+1-k} - \tilde{\pi}_k\tilde{\pi}_{i+j+1-k}];$$

$$N_0 = \sum_{m=0}^{n-2}(-1)^m x^{(m)}(0)\sum(P_{2k}-\tilde{P}_{2k})x^{(2k-m-1)}(0)$$

and the internal sum is taken for k: $2(m+1) < 2k < (n+m+1)$; N_T is obtained from N_0 by replacing $x^{(m)}(0)$ with $x^{(m)}(T)$; quantities with a tilde refer to polynomial $P_1(z)P_1(-z)$; P_{2k} is the coefficient at z^{2k} in the polynomial $P_0(z)P_0(-z)$.

Note. The proposed method of determining F[x(t)] is obviously applicable to the case where the mean values of $\xi(t)$ in hypotheses H_0 and H_1 are other than zero. For this

purpose, in place of x(t) in the right hand sides of equations (7) and (8) it is necessary to insert $x(t) + m_1(t)$ and $x(t) + m_0(t)$, respectively, where $m_0(t) = \eta_0(t)$ and $m_1 = \eta_1(t)$. We may also find F[x(t)] when hypotheses H_0 and H_1 differ only in mean values (see reference (8). In all these cases it is required that the mean values be sufficiently smooth functions.

Without calculations, let us present the functional F[x(t)] for one more case. Let

$$L(t-s) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1+\lambda^2}{1+\lambda^4} e^{i\lambda(t-s)}\, d\lambda,$$

$$K(t-s) = e^{-|t-s|}.$$

In this case

$$F[x(t)] = \lim_{n\to\infty}\left|\frac{L_n}{K_n}\right|^{\frac{1}{2}} \exp\left\{-\int_0^T x^2(t)\,dt + \frac{5}{4}\int_0^T\int_0^T e^{-|t-s|}\,x(t)\,x(s)\,dt\,ds + \right.$$

$$+ \frac{1}{25e^T - e^{-T}}\left[\frac{5}{2}e^{-T}\int_0^T\int_0^T e^{s-t}\,x(s)\,x(t)\,ds\,dt - \right.$$

$$- \frac{25e^{-T}}{2}\int_0^T\int_0^T \operatorname{ch}(s+t)\,x(s)\,x(t)\,dt\,ds + 10\,x(0)\,x(T) - e^{-T}(x^2(0)+x^2(T)) +$$

$$\left.\left. + 5\int_0^T (5e^t - e^{-t})\,[x(0)\,x(T-t) + x(T)\,x(t)]\,dt\right]\right\}.$$

As is seen from the presented examples, the exponent in F[x(t)] contains, in addition to the integral terms "terminal" terms of the form $x^2(0)$, $x^2(T)$, etc. The presence of these end terms is explained by the fact that processes with rational spectral density are components of multi-dimensional Markov processes. For such processes the values of the sample function at the ends of segment [O, T] and within it do not, roughly speaking, contribute equally. This may be explained as follows. For the problem of detection of a signal against a noise background any information concerning the process to be observed and delaying with its behavior beyond the discussed interval [), T] would be of value to us (for example, if we knew all the values of x(t) with $t < 0$ and $t > T$, we could solve without error the problem of the presence of a signal). But the greatest information concerning the values of the process beyond the interval [O, T] is, of course, conveyed by the values at the ends of the interval. Hence, the terminal values of the function x(t) carry greater weight in the expressions for F[x(t)] than the values of x(t) with $0 < t < T$. Let it be noted that with an increase in T the terminal terms in F[x(t)] play a decreasingly important role.

In obtaining the expression for F[x(t)] in explicit form we may effectively solve not only the problem of testing the statistical hypotheses but also the problem of estimating the unknown parameters from the sample function. For example, we may determine the meximum likelihood estimates possissing known optimum properties (see reference [3], page 229). We shall not discuss the subject further here.

2. DETERMINING THE CHARACTERISTIC FUNCTION OF A RANDOM QUANTITY ln F[ξ(t)]

The functional F[x(t)] determined in section 1 tells us what would be the optimum processing of observed data for testing the two hypotheses on the random process. Also of interest is the problem of probabilities of error in using this optimum processing. The answer lies in determining the distribution of the random value F[ξ(t)] in both hypotheses. Since the optimum criterion consists in comparing the values of F[ξ(t)] with a certain threshold constant Π with subsequent acceptance assumption of hypotheses H_0 and H_1 depending on whether F[ξ(t)] is greater than or less than Π, then it is easy to show that the use of any nomotonic function of F[ξ(t)] with appropriate change in constant Π will lead to the probabilities of error. It is most convenient for us to investigate the distribution of the random quantity $\eta = \ln F[\xi(t)]$. Since there is no direct method of determining the distribution function of η, it is necessary to limit ourselves to determining the characteristic function

of random quantity η. We will show how the characteristic function of η may be determined in the case where $\overline{\eta_0(t)} = \overline{\eta_1(t)} = 0$. Subsequent determination of the error probabilities may evidently be made by numerical methods alone. We will note that the statistical moments of η may be found either by expanding its characteristic function in a series or directly by using the explicit form of $F[\xi(t)]$.

It is self-evident that throughout this section we assume the fulfillment of condition (3). Otherwise, the error probabilities would be equal to zero. We will note that the method of determining the characteristic function of η has much in common with the method of determining the characteristic function of random quantity $\int_0^T \xi^2(t)dt$ (see reference [9]).

As the coordinates of the process $\xi(t)$ let us take the Fourier coefficients for the system of functions $\{\varphi_n(t)\}$:

$$\zeta_n = \int_0^T \xi(t)\,\varphi_n(t)\,dt,$$

where $\varphi_n(t)$ are the eigenfunctions of the integral equation

$$\int_0^T K(t-s)\,\varphi(s)\,ds = \lambda \int_0^T L(t-s)\,\varphi(s)\,ds, \quad 0 \leqslant t \leqslant T. \tag{12}$$

We will not discuss the problem of the "completeness" of the system $\{\varphi_n(t)\}$, for it is more our purpose to describe the methods of determining the limit of the likelihood ratio and the characteristic function $\ln F[\xi(t)]$ than to investigate the limits of applicability of these methods and to provide a rigid basis for them. In many specific cases, particularly in the example discussed at the end of this section, we are immediately aware that the system $\{\varphi_n(t)\}$ is such that the coordinates of $\{\zeta_n\}$ contains all the information concerning $\xi(t)$ at segment $[0, T]$.

Let us normalize $\varphi_k(t)$ so that

$$\int_0^T\int_0^T K(t-s)\,\varphi_k(t)\,\varphi_k(s)\,dt\,ds = 1. \tag{13}$$

By methods similar to those used in linear algebra for the simultaneous reduction of two quadratic forms to diagonal form (e.g., see reference [14], page 252) it is not difficult to show that in any choice of functions of $\{\varphi_n(t)\}$ the random quantities ζ_n will be independent in pairs both in hypothesis H_0 and in hypothesis H_1, that is,

$$\langle \zeta_m \zeta_n \rangle_{H_0} = \int_0^T\int_0^T K(t-s)\,\varphi_m(t)\,\varphi_n(s)\,dt\,ds = 0, \quad m \neq n,$$

$$\langle \zeta_m \zeta_n \rangle_{H_1} = \int_0^T\int_0^T L(t-s)\,\varphi_m(t)\,\varphi_n(s)\,dt\,ds = 0 \quad m \neq n$$

(the sign $\langle\rangle_{Hi}$ represents averaging for hypothesis H_i, $i = 0, 1$).

Due to condition (13) the dispersion of ζ_n when hypothesis H_0 holds is equal to unity and when hypothesis H_1 holds it is equal to $1/\lambda_n$.

The joint probability densities of quantities $\zeta_1 \ldots \zeta_n$ corresponding to hypotheses H_0 and H_1 will have the form

$$p_0(y_1 \ldots y_n) = (2\pi)^{-\frac{n}{2}} \exp\left\{-\tfrac{1}{2}\sum_{k=1}^{n} y_k^2\right\},$$

$$p_1(y_1 \ldots y_n) = (2\pi)^{-\frac{n}{2}} \left(\prod_{k=1}^{n} \lambda_k\right)^{\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\sum_{k=1}^{n} \lambda_k y_k^2\right\}.$$

Now let us insert into the likelihood ratio $p_0(y_1 \ldots y_n)/p_1(y_1 \ldots y_n)$ the values

$$x_i = \int_0^T x(t)\,\varphi_i(t)\,dt$$

and proceed to the limit:

$$F[x(t)] = \left(\prod_{k=1}^{\infty} \lambda_k\right)^{-\frac{1}{2}} \exp\left\{\frac{1}{2}\sum_{k=1}^{\infty} x_k^2(\lambda_k - 1)\right\}.$$

We are interested in the characteristic function of the random quantity

$$\eta = \ln F[\xi(t)] = \frac{1}{2}\sum_{k=1}^{\infty} \zeta_k^2(\lambda_k - 1) - \frac{1}{2}\ln\left(\prod_{1}^{\infty} \lambda_k\right).$$

For the hypothesis H_0 let us determine the characteristic function of η, which we shall designate as $\chi_0(t)$:

$$\chi_0(t) = \exp\left\{-\frac{it}{2}\ln\prod_{k=1}^{\infty}\lambda_k\right\}\left\langle \exp\left\{\frac{it}{2}\sum_{k=1}^{\infty}\zeta_k^2(\lambda_k - 1)\right\}\right\rangle_{H_0}.$$

Since ζ_k are independent,

$$\chi_0(t) = \exp\left\{-\frac{it}{2}\ln\prod_{k=1}^{\infty}\lambda_k\right\}\prod_{k=1}^{\infty}\left\langle \exp\left\{\frac{it\zeta_k^2}{2}(\lambda_k - 1)\right\}\right\rangle_{H_0}.$$

It is then easily shown that

$$\left\langle \exp\left\{\frac{it\zeta_k^2}{2}(\lambda_k - 1)\right\}\right\rangle_{H_0} = \frac{1}{\sqrt{1 - it(\lambda_k - 1)}},$$

wherein the argument of the derived function is so chosen that with t = 0 the function reverts to +1. Whence

$$\chi_0(t) = \frac{\exp\left\{-\frac{it}{2}\ln\prod_{k=1}^{\infty}\lambda_k\right\}}{\sqrt{\prod_{k=1}^{\infty}[1 - it(\lambda_k - 1)]}} = \frac{\exp\left\{-\frac{it}{2}\ln\prod_{k=1}^{\infty}\left(1 + \frac{1}{v_k}\right)\right\}}{\sqrt{\prod_{k=1}^{\infty}\left(1 - \frac{it}{v_k}\right)}},$$

where $v_k = 1/(\lambda - 1)$. Our problem now consists in obtaining the infinite product $\prod_{k=1}^{\infty}\left(1 - \frac{it}{v_k}\right)$ in closed form. This is done in quite the same manner as in determining the amount of information in one Gaussian process relative to another in reference [10]. First it is established that v_k are zeroes of a certain whole function D(v) of exponential type and of an order less than unity, which may be written explicitly. According to the Hadamard theorem (see reference [11]) for a whole function D(v) of exponential type and order less than unity there occurs expansion in coefficients

$$D(v) = D(0)\prod_{k=1}^{\infty}\left(1 - \frac{v}{v_k}\right), \text{ где } v_k \text{ — нули } D(v).$$

from which we find the infinite products we required

$$\prod_{k=1}^{\infty}\left(1 - \frac{it}{v_k}\right) = \frac{D(it)}{D(0)}, \quad \prod_{k=1}^{\infty}\left(1 + \frac{1}{v_k}\right) = \frac{D(-1)}{D(0)}.$$

Let us illustrate determination of $\chi_0(t)$ for the case where

$$K(t - s) = \frac{1}{\alpha}\exp\{-\alpha|t - s|\},$$

$$L(t - s) = \frac{1}{\beta}\exp\{-\beta|t - s|\}.$$

We now seek the solution of (12) in the form

$$\varphi(t) = \tilde{\varphi}(t) + a\delta(t) + b\delta(T - t).$$

Applying to both sides of the equality

$$\frac{1}{\pi}\int_0^T \widetilde{\varphi}(s)\,ds \int_{-\infty}^{\infty} \frac{e^{i\lambda(t-s)}}{\alpha^2+\lambda^2}\,d\lambda + \frac{a}{\alpha}e^{-\alpha t} + \frac{b}{\alpha}e^{-\alpha(T-t)} =$$

$$= \frac{1+\frac{1}{v}}{\pi}\int_0^T \widetilde{\varphi}(s)\,ds \int_{-\infty}^{\infty} \frac{e^{i\lambda(t-s)}}{\beta^2+\lambda^2}\,d\lambda + \left(1+\frac{1}{v}\right)\frac{a}{\beta}e^{-\beta t} + \left(1+\frac{1}{v}\right)\frac{b}{\beta}e^{-\beta(T-t)} \tag{15}$$

the operator $\left(-\frac{d^2}{dt^2}+\alpha^2\right)\left(-\frac{d^2}{dt^2}+\beta^2\right)$, we obtain the differential equation

$$\widetilde{\varphi}''(t) - [\alpha^2 + v(\alpha^2-\beta^2)]\,\widetilde{\varphi}(t) = 0,\ 0<t<T.$$

Whence

$$\widetilde{\varphi}(t) = C_1 e^{\Lambda t} + C_2 e^{-\Lambda t},$$

where

$$\Lambda = \sqrt{\alpha^2 + v(\alpha^2-\beta^2)}. \tag{16}$$

Let us apply to (15) the operator $\frac{d^r}{dt^r}\left(\frac{d}{dt}+\alpha\right)\left(\frac{d}{dt}+\beta\right)$, r = 0, 1 and assume t = T. We obtain the two equations

$$\begin{aligned} &\widetilde{\varphi}(T) - b(\alpha+\beta) = 0,\\ &\widetilde{\varphi}'(T) + (\alpha+\beta)\,\widetilde{\varphi}(T) + b\{(\alpha^2-\beta^2)v - \beta(\alpha+\beta)\} = 0. \end{aligned} \tag{17}$$

In similar fashion we obtain another two equations at point t = 0.

$$\begin{aligned} &\widetilde{\varphi}(0) - a(\alpha+\beta) = 0,\\ &\widetilde{\varphi}'(0) - (\alpha+\beta)\,\widetilde{\varphi}(0) - a\{(\alpha^2-\beta^2)v - \beta(\alpha+\beta)\} = 0. \end{aligned} \tag{18}$$

Inserting into Equations (17) and (18) $\widetilde{\varphi}(0)$, $\widetilde{\varphi}'(0)$, $\widetilde{\varphi}(T)$ and $\widetilde{\varphi}'(T)$ in the form of (16), we obtain for C_1, C_2, a, b, the equations

$$\begin{aligned} &C_1 e^{\Lambda T} + C_2 e^{-\Lambda T} - b(\alpha+\beta) = 0,\\ &C_1 e^{\Lambda T}[\Lambda + (\alpha+\beta)] + C_2 e^{-\Lambda T}[-\Lambda + (\alpha+\beta)] + b\{(\alpha^2-\beta^2)v - \beta(\alpha+\beta)\} = 0,\\ &C_1 + C_2 - a(\alpha+\beta) = 0,\\ &C_1[\Lambda - (\alpha+\beta)] - C_2[\Lambda + (\alpha+\beta)] - a\{(\alpha^2-\beta^2)v - \beta(\alpha+\beta)\} = 0. \end{aligned}$$

The determinant of the system is

$$D(v) = -(\alpha+\beta^2)\{e^{\Lambda T}[\Lambda + (\alpha-\beta)v + \alpha]^2 - e^{-\Lambda T}[\Lambda - (\alpha-\beta)v - \alpha]^2\}.$$

Function D(v) contains one excess root $v = \alpha^2/(\beta^2 - \alpha^2)$ which is not an eigenvalue of Equation (15). If $v = \alpha^2/(\beta^2 - \alpha^2)$, then the general solution of $\widetilde{\varphi}(t)$ is written not in the form of (16) but in the form of a second-degree polynomial. Dividing D(v) by Λ, we eliminate this root. Finally, for $\chi_0(t)$ we have

$$\chi_0(t) = \sqrt{\frac{D(0)\,\Lambda(it)}{D(it)\,\Lambda(0)}}\exp\left\{-\frac{it}{2}\ln\frac{D(-1)\,\Lambda(0)}{D(0)\,\Lambda(-1)}\right\} =$$

in which

$$= \frac{2\sqrt{\alpha\Lambda}\left(\frac{\alpha}{\beta}\right)^{\frac{it}{2}}\exp\left\{\frac{it(\alpha-\beta)+\alpha}{2}T\right\}}{\sqrt{e^{\Lambda T}[\Lambda + it(\alpha-\beta)+\alpha]^2 - e^{-\Lambda T}[\Lambda - it(\alpha-\beta)-\alpha]^2}}, \tag{19}$$

In a similar manner it can also be shown that in the case of rational spectral densities the values of v_k in expression (14) are zeroes of a function satisfying the Hadamard theorem. This function, as in the example discussed by us, may be written explicitly in the form of a determinant and, consequently, we may obtain the characteristic function ln F[ξ(t)] in closed form. Solution of equation (12 must be written in the form

$$\Lambda = \sqrt{\alpha^2 + it(\alpha^2-\beta^2)},\ \Lambda(0) = \alpha,\ \Lambda(-1) = \beta.$$

$$\varphi(t) = \widetilde{\varphi}(t) + \sum_{k=0}^{l} a_k \delta^{(k)}(t) + \sum_{k=0}^{l-1} b_k \delta^{(k)}(t-T),$$

where l is the difference in degree of $P_0(i\lambda)$ and $Q_0(i\lambda)$ and $\widetilde{\varphi}\, t$ is the solution of the equation

$$P_1\left(\frac{d}{dt}\right) P_1\left(-\frac{d}{dt}\right) Q_0\left(\frac{d}{dt}\right) Q_0\left(-\frac{d}{dt}\right) \widetilde{\varphi}(t) =$$
$$= \lambda P_0\left(\frac{d}{dt}\right) P_0\left(-\frac{d}{dt}\right) Q_1\left(\frac{d}{dt}\right) Q_1\left(-\frac{d}{dt}\right) \widetilde{\varphi}(t).$$

Let us make one remark concerning determination of the distribution of ln F[ξ(t)]. We may conclude from a few general observations that with T → ∞ ln F[ξ(t)] will be a normal asymptotic quantity.

In particular, for the example discussed by us it is not difficult to show that

$$\exp\left\{-it\frac{a_k}{b_k}\right\} \chi_k\left(\frac{t}{b_k}\right) \to \exp\left\{-\frac{t^2}{2}\right\}$$

with T → ∞, k = 0.1. Here $a_0 = \frac{(\alpha-\beta)^2 T}{4\alpha}$;

$$a_1 = -\frac{(\alpha-\beta)^2 T}{4\beta}; \quad b_0 = \frac{|\beta^2-\alpha^2|\sqrt{T}}{(2\alpha)^{3/2}}; \quad b_1 = \frac{|\beta^2-\alpha^2|\sqrt{T}}{(2\beta)^{3/2}}.$$

The parameters of the normal limiting distribution may be found by calculating the mean and the dispersion of ln F[ξ(t)], whereupon an explicit expression is obtained for it, or by expanding into a Taylor series the characteristic function ln F[ξ(t)].

We shall not discuss in this report the problem of the errors which arise in replacing the true distribution of ln F[ξ(t)] by the asymptotic distribution. In order to evaluate ;this error it is obviously best to use the explicit form of the characteristic function ln F[ξ(t)].

3. CONCERNING A MORE REALISTIC STATEMENT OF THE PROBLEM OF DETECTION OF A SIGNAL AGAINST A NOISE BACKGROUND

As already stated in section 1, in satisfying condition (2) we may know without error which of the hypotheses holds true. Let us discuss the methods by which this may be done. Let $K(t-s) = \frac{a}{\alpha} e^{-\alpha|t-s|}$, $L(t-s) = \frac{b}{\alpha} e^{-\alpha|t-s|}$ and a > b. In this case condition (2) is satisfied, since

$$\lim_{\lambda\to\infty} \left| \frac{Q_0(i\lambda)\, P_1(i\lambda)}{P_0(i\lambda)\, Q_1(i\lambda)} \right|^2 = \frac{a}{b} > 1.$$

Let us present, after Slepian [4], the sum

$$Y_n[x(t)] = \sum_{j=1}^{n} \left\{ x\left(\frac{Tj}{n}\right) - x\left(\frac{T(j-1)}{n}\right) \right\}^2 \tag{20}$$

and let n → ∞. Then, as shown in reference [4], for sample functions of the process $\eta_0(t)$ we have the equality

$$\lim_{n\to\infty} Y_n[x(t)] = 2aT,$$

and for all sample functions of the process $\eta_1(t)$ we have the equality

$$\lim_{n\to\infty} Y_n[x(t)] = 2bT.$$

Since a > b, at any T we may know without error which of the hypotheses is true. Let it be noted that in the case of fulfilment of condition (2) we may propose different variants of errorless testing of hypotheses H_0 and H_1. However, in practice not one of these variants can be used, for the values of the function and spectral densities of processes $\eta_0(t)$ and $\eta_1(t)$ are known with only limited accuracy.

The paradox of errorless testing of the hypotheses on condition (2) is based on use of the difference in the high-frequency portion of the spectra. In addition, we use the local

properties of sample functions of the process, which are essentially associated with the asymptotic behavior of the spectrum and not with its behavior in the terminal region of frequencies. In these cases errorless testing of the hypotheses is actually based on frequencies which are so high as to have no physical meaning.

In order that the problem of testing hypotheses concerning the sample function of a random process would be more realistic, L. A. Vaynshteyn proposed that it be assumed that, in addition to processes $\eta_0(t)$ or $\eta_1(t)$, there is present on each occasion white noise with the same low spectral density. With the addition of white noise condition (3) is fulfilled and the paradoxes of errorless testing of the hypotheses vanish. White noise does not permit the use of information which carries infinitely high frequencies, for these are merely a mathematical abstraction.

Thus, we shall consider that the correlation functions corresponding to hypotheses H_0 and H_1 have the form

$$\widetilde{K}(t-s) = N\delta(t-s) + \frac{a}{\alpha}e^{-\alpha|t-s|}, \quad \widetilde{L}(t-s) = N\delta(t-s) + \frac{b}{\alpha}e^{-\alpha|t-s|},$$

where N is a small number; $\delta(t-s)$ is the Dirac function, which is the correlation function of white noise. In the manner described in section 1 we find $F_N[x(t)]$:

$$F_N[x(t)] = \lim_{n\to\infty}\left|\frac{\widetilde{L}_n}{\widetilde{K}_n}\right|^{\frac{1}{2}} \exp\left\{\frac{1}{2N^2}\int_0^T\int_0^T\left[\frac{a}{\alpha_1}e^{-\alpha_1|t-s|} - \frac{b}{\beta_1}e^{-\beta_1|t-s|}\right]\times\right.$$

$$\times x(t)\,x(s)\,dt\,ds + \frac{1}{2N}\int_0^T\int_0^T\left[\frac{e^{-2\alpha_1 T}(\alpha_1-\alpha)^3\operatorname{ch}\alpha_1(t-s)}{\alpha_1(\alpha+\alpha_1)\left[1-\left(\frac{\alpha_1-\alpha}{\alpha_1+\alpha}\right)^2 e^{-2\alpha_1 T}\right]} -\right.$$

$$-\left.\frac{e^{-2\beta_1 T}(\beta_1-\beta)^3\operatorname{ch}\beta_1(t-s)}{\beta_1(\beta+\beta_1)\left[1+\left(\frac{\beta_1-\beta}{\beta_1+\beta}\right)^2 e^{-2\beta_1 T}\right]}\right]x(t)\,x(s)\,dt\,ds +$$

$$+\frac{1}{2N}\int_0^T\int_0^T\left[\frac{e^{-\alpha_1 T}(\alpha_1-\alpha)^2\operatorname{ch}\alpha_1(s+t-T)}{\alpha_1\left[1-\left(\frac{\alpha_1-\alpha}{\alpha_1+\alpha}\right)^2 e^{-2\alpha_1 T}\right]} -\right.$$

$$-\left.\left.\frac{e^{-\beta_1 T}(\beta_1-\beta)^2\operatorname{ch}\beta_1(s+t-T)}{\beta_1\left[1-\left(\frac{\beta_1-\beta}{\beta_1+\beta}\right)^2 e^{-2\beta_1 T}\right]}\right]x(t)\,x(s)\,dt\,ds\right\}, \tag{21}$$

where

$$\alpha_1 = \sqrt{\frac{2a}{N}+\alpha^2} \text{ and } \beta_1 = \sqrt{\frac{2b}{N}+\alpha^2}.$$

Upon fulfillment of the conditions

$$2\frac{a}{\alpha} \gg N\alpha, \quad 2\frac{b}{\alpha} \gg N\alpha, \tag{22}$$

$$\alpha T \gg 1 \tag{23}$$

the expression obtained for $F_N[x(t)]$ may be considerably simplified. Physically, the condition $2\frac{a}{\alpha} \gg N\alpha$ expresses the natural requirement that the added white noise have small power in the frequency band from 0 to α (that is, in the band where the spectral density of $\alpha/\pi(\alpha^2+\lambda^2)$, which is equal to a/α. Condition $\alpha T \gg 1$ means that in the observation interval $[0, T]$ there are many intervals of correlation of the process with the correlation function $\frac{a}{\alpha}e^{-\alpha|t-s|}$.

Conditions (22) and (23) ensure smallness of the second and third terms in the exponent of (21) in comparison with the first term. Thus, for such small N and for such T that conditions (22) and (23) are fulfilled, we shall have

$$F_N[x(t)] \simeq \lim\left|\frac{\widetilde{L}_n}{\widetilde{K}_n}\right|^{\frac{1}{2}} \exp\left\{\frac{1}{2N^2}\int_0^T\int_0^T\left[\frac{a}{\alpha_1}e^{-\alpha_1|t-s|} - \frac{b}{\beta_1}e^{-\beta_1|t-s|}\right]x(t)\,x(s)\,dt\,ds\right\}.$$

It may be shown that for any continuous function x(t) with $N \to 0$, with an accuracy of the order of infinitely small terms of higher order in N there exists an approximate equality

$$F_N[x(t)] \simeq \lim \left| \frac{\widetilde{L}_n}{\widetilde{K}_n} \right|^{\frac{1}{2}} \exp\left\{ \frac{C}{\sqrt{N}} \sum_{j=1}^{\left[\frac{T}{2\tau_0}\right]} [x(2\tau_0 j) - x(2\tau_0 (j-1))]^2 \right\}, \quad (24)$$

where $\tau_0 = \frac{\ln \frac{a\beta_1}{b\alpha_1}}{\alpha_1 - \beta_1} \simeq \frac{\ln \frac{a}{b}}{2\sqrt{2}(\sqrt{a} - \sqrt{b})}\sqrt{N}$; C is a certain constant independent of N. The exponential in (24) contains a sum in the form of (2). However, due to the fact that the square of the white noise has no meaning, it is not possible to insert into (24) x(t) containing white noise. But, if it is assumed that with "reduction of the white noise (for example, so that the spectral density becomes equal to $f(\lambda) = \begin{cases} N, & |\lambda| < N_0 \\ \frac{NN_0^2}{\lambda^2}, & |\lambda| \geqslant N_0 \end{cases}$ for sufficiently large N_0 the optimum method of testing the hypotheses undergoes little change, then expression (24) will provide an approximately optimum method of processing the data for testing the hypotheses. For the "reduced" white noise, squaring then has meaning. It is interesting to note that the method proposed by Slepian (which upon satisfying condition (2), permits errorless testing of the hypotheses) remains optimum upon the addition of a small amount of reduced white noise if the interval of the sum in (20) does tend toward zero and permits selection in accordance with the spectral density of the noise. With a decrease in noise density N the interval $2\tau_0$ decreases in proportion of $\sqrt{N}$. The characteristic functions of $F_n[\xi(t)]$, calculated by the method proposed in section 2, will have the form

$$\chi_0^{(N)}(t) = \exp\left\{ -\frac{it}{2} \ln \frac{D(-1)}{D(0)} \right\} \sqrt{\frac{D(0)}{D(it)}},$$

where

$$D(v) = \frac{e^{k(v)T}[k(v) - \alpha]^2 - e^{-k(v)T}[k(v) + \alpha]^2}{k(v)};$$

$$k(v) = \sqrt{\alpha^2 + \frac{2b}{N} + \frac{2v}{N}(b - a)};$$

$\chi_1^{(N)}(t)$ is obtained from $\chi_0^{(N)}(t)$ by replacing a_1, b, t with b_1, a, -t, respectively.

Designating the Fourier transforms of $\chi_0^{(N)}(t)$ and $\chi_1^{(N)}(t)$ by $p_0^{(N)}(x)$ and $p_1^{(N)}(x)$, we may write the error probabilities of the first and second kind in this more realistically stated problem in the form $\int_{-\infty}^{\Pi} p_0^{(N)}(x)\,dx$ and $\int_{\Pi}^{\infty} p_1^{(N)}(x)\,dx$. It may be shown that with fixed T and $N \to 0$ these probabilities also tend toward zero. Conversely, with fixed N and $T \to 0$ they increase to certain limit values.

CONCLUSION

Gaussian stationary processes with rational spectral density have been investigated for a finite segment of time. Expressions have been derived which show what should be the optimum processing of the sample function of a random process for the testing of two competing hypotheses concerning this process. A method has been given for determining in closed form the characteristic function of the limiting likelihood ratio. The problem of testing two hypotheses concerning the sample function of the process has been discussed with an approach that renders the problem more realistic.

The author expresses his thanks to R. L. Dobrushin, who directed this work, for his numerous helpful suggestions and also to M. A. Isakovich for his valuable comments on the work.

REFERENCES

1. Peterson, Birdsall, Fox, Theory of signal detection; a collection of translations under the editorial supervision of A. A. Kharkevich; Teoriya informatsii i yeye prilozheniya. [Information Theory and Its Applications], GIFML, 1959.
2. D. Middleton, On the detection of stochastic signals in additive normal noise; Part I; IRE Trans., 1957, IT-3, 2, 86.
3. U. Grenander, Stochastic processes and statistical inference, Ark. Mat., 1950, 1, 195.
4. D. Slepian, Some comments on the detection of Gaussian signals in Gaussian noise, IRE trans., 1958, IT-4, 2, 65.

5. A.N. Kolmogorov, Osnovnyye ponyatiya teorii veroyatnostey, [Basic Concepts of Probability Theory], ONTI, 1936.
6. L. Zadeh, J. Ragazzini. Optimum filters for the detection of signals in noise, Proc. IRE, 1952, 40, 1223.
7. E. Reich, P. Swerling, Detection of sine wave in Gaussian noise, J. Appl Phys., 1953, 24, 3, 289.
8. R. Davis, The detectability of random signals in the presence of noise, IRE Trans., 1954, PGIT-3, 52.
9. D. Slepian, Fluctuation of random noise power, Bell System Tech. J., 1958, 37, 1, 163.
10. I.M. Gel'fand, A.M. Yaglom, On calculation of the amount of information concerning a random function, Uspekhi matem, nauk, 1957, 12, 1, 3.
11. A.I. Markushevich, Teoriya analiticheskikh funktsiy [Theory of Analytical Functions], GITTL, 1950.
12. W. Davenport, W. Root, An Introduction to the Theory of Random Signals and Noise, 1958, N.Y.
13. C. Striebel, Densities for stochastic processes, Ann. Math. Stat., 1959, 30, 2, 559.
14. G.Ye. Shilov, Vvedeniye v teoriyu lineynykh prostranstv, [Introduction to the Theory of Linear Amplitude], GITTL, 1952.

Machine Mathematics Department
of M.B. Lomonosov Moscow State University
Chair of Probability Theory

Submitted to the editors, 3 December 1959

INFLUENCE OF STORAGE-DEVICE PARAMETERS ON OPERATIONAL EFFICIENCY

Yu. S. Lezin

The signal-to-noise ratio at the output of one and several storage devices with delayed feedback is calculated. The influence of passband and feedback factor on the signal-to-noise ratio of these devices is analyzed.

INTRODUCTION

The operational efficiency of storage devices designed to increase the signal-to-noise ratio must be evaluated by the value of this ratio at the output for given signal power and intensity of the noise-power spectrum at the input or by the gain in signal-to-noise ratio provided by these devices.

A storage device with delayed feedback [1], the block diagram of which is shown in Figure 1 [2], is characterized by two parameters: the filter passband ΔF and the feedback factor m (at sufficiently low frequencies). Let us discuss the influence of these parameters on the operational efficiency of the storage device. We shall assume that at the input of the storage device there is applied a train of rectangular pulse signals of identical amplitude U_1 and identical duration τ, the number N of which is so great that

plus white noise with normal distribution. In addition, the filter passband ΔF is assumed to be much greater than the pulse-repetition frequency $F = 1/T$.

1. A SINGLE STORAGE DEVICE WITH DELAYED FEEDBACK

1. Basic Relationships. As shown in reference [3], the maximum pulse signal amplitude at the output of a single storage device is

$$U_2 = \frac{U_1}{1-m}[1 - e^{-2\pi\tau\Delta F(1-m)}]. \tag{1}$$

In reference [2] it was established that the noise at the output of the given storage device has a power

$$\sigma_2^2 = \frac{\pi^2 a \Delta F}{\sqrt{1-m^2}}, \tag{2}$$

where a is the intensity of the noise-power spectrum at the input. Hence the signal-to-noise power ratio at the output is

$$A = A_0 \frac{\sqrt{1-m^2}}{(1-m)^2}\frac{1}{\pi b}[1 - e^{-2\pi b(1-m)}]^2, \tag{3}$$

where $A_0 = U_1^2 \tau/\pi$ a is the signal -to-noise ratio at the output of an optimal filter for a single pulse signal [4]: b = $\Delta F\tau$ is the product of the filter passband and the pulse duration (it may be referred to as the "dimensionless" passband). Consequently, the gain in signal-to-noise ratio provided by such a storage device, in comparison with an optimal filter for a single pulse signal is

$$B_1 = \frac{\sqrt{1-m^2}}{(1-m)^2}\frac{1}{\pi b}[1 - e^{-2\pi b(1-m)}]^2. \tag{4}$$

2. Influence of Passband. Investigating (4) at the maximum with respect to b, let us determine the optimum passband

$$b_1 = \Delta F_1 \tau = \frac{0{,}200}{1-m}. \tag{5}$$

It increases sharply as the feedback factor approaches unity (Curve 1 in Figure 1)

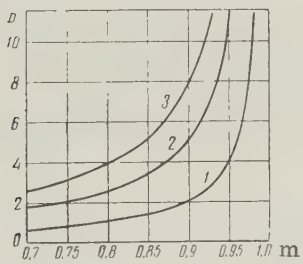


Figure 1. Dependence of optimum passband on feedback factor for one, two, and three storage devices.

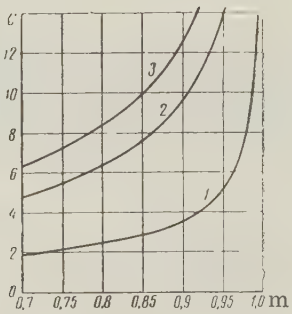


Figure 2. Dependency of maximum gain on feedback factor for one, two, and three storage devices.

With a given feedback factor, the maximum value of gain in signal-to-noise ratio due to the use of a storage device is achieved with optimum bandwidth and is equal to

$$C_1 = B_1(b = b_1) = 0{,}815\sqrt{\frac{1+m}{1-m}}. \tag{6}$$

This gain increases rapidly as the feedback factor approaches unity (curve 1 in Figure 2) and at m = 0.9 is 3.5. High gain is obtained by increasing the feedback factor, but in practice this may lead to instability in operation of the storage device.

The amount of loss in signal-to-noise ratio due to nonoptimality of passband is defined by the ratio B_1/C_1, which in accordance with (4) and (6) is equal to

$$\frac{B_1}{C_1} = 1.954 \frac{(1 - e^{-1.256z})^2}{z}, \tag{7}$$

where $z = b/b_1 = \Delta F/\Delta F_1$ is the ratio of the passband to its optimum value. From the dependence of loss in signal-to-noise ratio on z (curve 1 in Figure 3) it follows that the optimum value of passband is not critical.

3. Influence of Feedback Factor. Investigating (4) at the maximum with respect to m, let us determine the optimum value of feedback factor as a function of dimensionless passband (curve 1 in Figure 4). With an increase in passband the optimum value of feedback factor at first rises rapidly and then slowly, approaching unity. With b = 1 it is 0.917 and with b = 2 it is already 0.957. The maximum gain in signal-to-noise increases with passband (Curve 1 in Figure 5) and is 3.06 with b = 1 and 4.37 with b = 2.

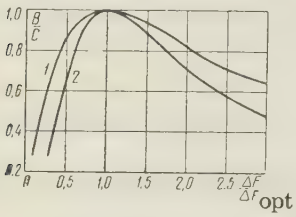


Figure 3. Ratio of gain to maximum as a function of the ratio of passband to optimum passband for one and two storage devices.

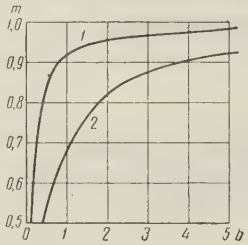


Figure 4. Optimum value of feedback factor as a function of passband for one and two storage devices.

Figure 6 shows a family of curves for gain in signal-to-noise power ratio as a function of feedback factor with various values of passband. From an examination of these curves it follows that deviation of the feedback factor from optimum value leads to a slight decrease in signal-to-noise ratio with b < 0.5 and to a sharp decrease in this ratio with b > 1, wherein the degree of this decrease rises with an increase in passband.

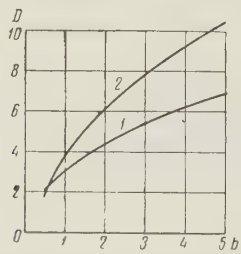


Figure 5. Dependence of maximum gain on passband for one and two storage devices.

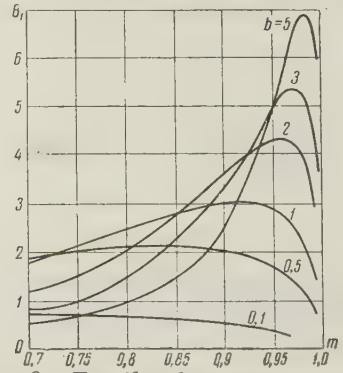


Figure 6. Family of curves for gain in signal-to noise ratio as a function of feedback factor for various values of passband in the case of a single storage device.

4. Gain Provided by a Storage Device in Comparison with a Low Pass Filter. In order to determine storage efficiency in comparison with conventional frequency filtering let us determine the gain provided by a storage device in comparison with a low pass filter, the transfer function of which has the form

$$K^*(\omega) = \frac{1}{1 + j\frac{\omega}{2\pi\Delta f}}, \tag{8}$$

where Δf is the passband of this filter.

Using (4) with m = 0 and also using (6), it is easily shown that in comparison with a low pass filter a storage device provides the following maximum gain in signal-to-noise ratio:

$$R_1 = \frac{C_1}{B_1(m=0; b=q)} = \frac{2{,}56q}{(1-e^{-2\pi q})^2}\sqrt{\frac{1+m}{1-m}}, \tag{9}$$

where $q = \Delta f\tau$ is the dimensionless passband of the low pass filter. In particular, with q=1

$$R_{11} = 2{,}57\sqrt{\frac{1+m}{1-m}} = 3{,}15\,C_1, \tag{10}$$

and with optimum passband q = 0.2 (which is easily determined from (5) assuming therein that m = 0 and replacing b with q)

$$R_{12} = \sqrt{\frac{1+m}{1-m}} = 1{,}227\,C_1. \tag{11}$$

With m = 0.9, this gain is 11.2 and 4.36 respectively.

Consequently, the use of a single storage device with delayed feedback yields, in comparison with a filter, a relatively large gain in signal-to-noise ratio.

2. SERIES-CONNECTED STORAGE DEVICES WITH DELAYED FEEDBACK

1. Two Storage Devices. Using the expressions for maximum pulse-signal amplitude and noise power at the output of two identical, series-connected storage devices (derived respectively, in references [3] and [2]), let us determine the gain in signal-to-noise ratio provided by two devices in comparison with an optimal filter for a single pulse signal:

$$B_2 = \frac{4}{\pi b}\,\frac{(1-m^2)^{3/2}}{(1-m)^4(2+m^2)}\{1-[1+2\pi b(1-m)]\,e^{-2\pi b(1-m)}\}^2. \tag{12}$$

The optimum passband in the case of two storage devices

$$b_2 = \Delta F_2\tau = \frac{0{,}512}{1-m} \tag{13}$$

increases as m approaches unity (curve 2 in figure 1) as in the case of a single storage device, but in distinction from the latter case it is 2.56 times greater.

With the optimum passband of two storage devices the gain in signal-to-noise ratio is maximum:

$$C_2 = 1{,}714\,\frac{(1+m)^2}{2+m^2}\left(\frac{1+m}{1-m}\right)^{1/2}. \tag{14}$$

It increases as the feedback factor approaches unity (curve 2 in Figure 2) and with m = 0.9 it is 9.56.

The additional maximum gain due to use of the second storage device is

$$G_2 = \frac{C_2}{C_1} = 2{,}10\,\frac{(1+m)^2}{2+m^2} \tag{15}$$

and with $0.7 < m < 1$

$$G_2 \simeq 1{,}6 + 1{,}2m.$$

This gain in the given interval of m rises monotonically from 2.44 to 2.80.

Thus, the maximum signal-to-noise ratio at the output of two storage devices is 2.5-2.8 times greater than at the output of a single storage device. However, the storage arrangement becomes considerably more complex and it is necessary to increase the passband by

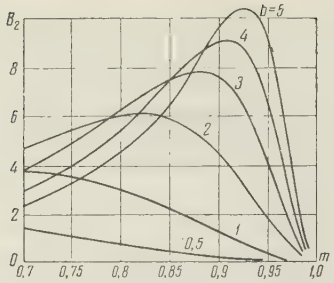


Figure 7. The family of curves for gain in signal-to-noise power ratio as a function of the feedback factor for various passband values with two storage devices.

2.56 times.

Examination of the curve for loss in signal-to-noise ratio due to nonoptimality of passband as a function of the ratio of the passband to optimum (curve 2 in Figure 3) shows that in the case of two storage devices the optimum passband proves to be more critical than in the case of a single storage device.

With two storage devices the optimum feedback factor increases with an increase in passband (curve 2 in Figure 4), but its value for the same passband is considerably smaller than in the case of a single storage device. Consequently, in using two storage devices, in order to ensure high gain in the signal-to-noise ratio (with a given passband) there is no need to bring the feedback factor to a value close to unity.

The maximum gain in signal-to-noise ratio with two storage devices increases with passband more rapidly (curve 2 in Figure 5) than with one storage device. The additional gain due to the second storage device rises slightly with an increase in passband and with b = 3 is approximately 1.5.

Figure 7 shows the family of curves for gain in signal-to-noise power ratio as a function of the feedback factor at six values of passband.. From an examination of this family of curves it is seen that the maxima of the curves at b > 1 are flatter than with a single storage device (see Figure 6). Consequently, the optimum value of feedback factor in this case is less than the critical value.

2. Three Storage Devices. By use of the results obtained in references [3] and [2], respectively, for the maximum pulse signal amplitude and noise power at the output of three identical series-connected storage devices, let us determine the gain provided by three storage devices in comparison with the gain of an optimal filter for a single pulse signal:

$$B_3 = \frac{64}{3\pi b} \frac{(1-m^2)^{3/2}}{(1-m)^6(8+24m^2+3m^4)} \{1 - [1 + 2\pi b(1-m) + 2\pi^2 b^2 (1-m)^2] e^{-2\pi b(1-m)}\}^2. \quad (16)$$

With three storage devices the optimum passband increases with the feedback factor (Curve 3 in Figure 1), but its value is 3.92 and 1.53 times greater than with one and two storage devices, respectively. With three storage devices the dependence of maximum gain in signal-to-noise ratio on the feedback factor

$$C_3 = \frac{6.56(1+m)}{8+24m^2+3m} \left(\frac{1+m}{1-m}\right)^{1/2} \quad (17)$$

lies somewhat higher (curve 3 in Figure 2) than in the case of two storage devices.

The use of third storage device gives an additional gain of

$$G_3 = \frac{C_3}{C_2} = 3.8 \frac{(1+m)^2(2+m^2)}{8+24m^2+3m}, \quad (18)$$

Thus, use of a third storage device yields a relatively small gain in signal-to-noise ratio. Since this also introduces considerable complication of the storage arrangement, the use of a third storage device has little to offer.

CONCLUSION

It has been established above that filter passbands and the feedback factors of storage devices with delayed feedback have optimum values at which the filters and storage devices operate most efficiently.

The optimum value of passband rises sharply as the feedback factor approaches unity; it is only slightly critical with a single storage device and is more critical with two storage devices. An increase in the number of storage devices to two and three leads to an increase

in optimum passband by 2.56 and 3.92 times, respectively.

The maximum gain in signal-to-noise using storage devices as compared with the use of an optimal filter for a single pulse signal and the use of frequency filters increases with the feedback factor approaching unity and with an increase in the number of storage devices. At m = 0.9 the value of this gain in comparison with single-pulse optimal filtering is 3.5, 9.56 and 12.5 times for one, two and three storage devices, respectively.

The use of the second storage device provides an additional gain of 2.4 - 2.8 and the use of the third storage device provides an additional gain of 30%. Hence, the use of a second storage device in a number of cases will prove to be justified despite the considerable complication of the storage arrangement, whereas the use of a third storage device will not be so justified.

An additional advantage of the use of the second storage device is the lower critical optimum value of feedback factor, which, moreover, is lower than the optimum value of feedback factor in the case of a single storage device. However, there also arises the necessity for increasing the passband by 2.56 times.

Using the results obtained above and the results of any study on coherent detection of a single signal, it is easy to determine the threshold signals for coherent storage by means of devices with delayed feedback.

The signal-to-noise ratio at the output of a storage system may be further increased by several times if a narrowband filter is placed before the delayed-feedback storage devices; this subject, however, is beyond the scope of this report.

REFERENCES

1. W.D.White, A.E. Ruvin, Recent advances in the synthesis of comb filters, Convent, IRE, 1957, 5, part 2, 186.
2. Yu.S. Lezin, Storage of noise in devices with delayed feedback, Radiotekhnika i elektronika, 1961, 6, 2, 187.
3. L.A. Morugin, Storage of pulse signals in devices with delayed feedback, Radiotekhnika i Elektronika, 1960, 5, 12, 1885.
4. Yu. S. Lezin, Passage of signal and noise through an optimal filter, Tr. Gor'kovskogo politekhnicheskogo instituta im. A.A. Zhdanova, 1958, 14, 5, 45.

Submitted to the editors 28 March 1960
After revision 30 September 1960

RANGE FINDER WITH FREQUENCY MODULATION IN THE PRESENCE OF NOISE AND FLUCTUATIONS IN REFLECTED SIGNAL

G. P. Tartakovskiy

In this report formulas are derived for systematic errors and fluctuation errors in a frequency-modulated range finder which occur in the presence of receiver noise and fluctuations in the reflected signal. The dependence of these errors on various parameters is discussed and several recommendations are made.

INTRODUCTION

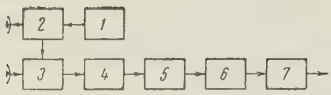


Figure 1. Block diagram of frequency modulated rangefinder: 1, modulator; 2, transmitter; 3, mixer; 4, amplifier and filter; 5, limiter; 6, pulse counter 7, d-c amplifier.

In CW radars and in aircraft altimeters the frequency-modulated rangefinder (Figure 1) is often used.

The transmitter radiates frequency-modulated oscillations which are also used as reference oscillations for the mixer. Delayed reflected oscillations reach the input of the mixer, with the result that at the output of the filter (possessing a transmission band in the low-frequency region) there appear oscillations such that the greater the number of zeros therein, the greater the delay of the reflected oscillations. (i.e., the measured range). The number of zeros during measurement time T is measured by a counter, which may consist for example, of a trigger-operated blocking oscillator and a pulse counter. The processes in a rangefinder of the described type have been analyzed in reference [1] where, in particular, there is shown the nonproportionality of a simplified treatment in accordance with which the rangefinder measures the frequency of oscillations at the filter output. In this report the received oscillations were analyzed as a deterministic process differing from the radiated oscillations only in the delay.

Actually, however, the reflected oscillations are always random due to unavoidable fluctuations. Experience showns that neglect of fluctuations, particularly under moving-target conditions, leads to considerable error in evaluating the accuracy of measurement devices. Moreover, reception of a radar signal is always accompanied by receiver noise, which leads to additional errors.

The present report investigates errors in a frequency-modulated rangefinder in the presence of noise and fluctuations in the reflected signal.

1. CORRELATION FUNCTION OF THE PROCESS AT THE FILTER OUTPUT OF A RANGEFINDER

The main signal may be represented in the form

$$U(t) = u_0 \cos[\omega_0 t + \psi(t)],$$

where $\psi(t)$ is the phase of the frequency-modulated oscillation. Henceforth in our analysis we shall assume for the sake of concreteness that

$$\psi(t) = \frac{\Delta\omega}{\Omega_M} \sin \Omega_M t, \tag{2}$$

where Δ_{ω} is the frequency deviation and Ω_M is the modulation frequency.

The reflection object may usually be represented as a large number of independent elementary reflectors subject to slow, random displacement. The result is that the reflected signal is a normal, narrow band random process with respect to ω_0.

It may be shown that in the presence of symmetrical (relative to ω_0) spectral density of fluctuations* and in the presence of the main signal (1) the correlation function of the reflected signal

$$R_{вхс}(t, \tau) = u^2 \cos[\omega_0 \tau + \psi(t - \tau_0) - \psi(t - \tau - \tau_0)]\, \rho(\tau), \tag{3}$$

where $\rho(\tau)$ is the envelope of the fluctuation correlation coefficient in the presence of the main signal in the form $\sin \omega_0 t$, and τ_0 is the delay of the reflected signal.

The receiver processes the additive mixture of signal with noise x(t) the correlation function of which is

$$R_x(t, \tau) = R_{вхс}(t, \tau) + R_{вхш}(\tau) = \\ = u^2 \cos[\omega_0 \tau + \psi(t - \tau_0) - \psi(t - \tau - \tau_0)]\, \rho(\tau) + N_0 \delta(\tau). \tag{4}$$

*By this is meant fluctuations of the reflected signal in the presence of a main signal in the form of a sinusoid of frequency ω_0.

It is assumed that the receiver noise is white with a two-sided spectral density N_0 and $\delta(\tau)$ is the delta-function.

Within the mixer is formed the product

$$y(t) = x(t)\,U(t). \tag{5}$$

The correlation function of the process y(t) is determined as

$$R_y(t,\ \tau) = R_{s,}(t,\ \tau) + R_{n,}(t,\ \tau). \tag{6}$$

In expression (6)

$$R_{s,}(t,\tau) = u^2 \cos[\omega_0\tau + \psi(t-\tau_0) - \psi(t-\tau-\tau_0)]\,\rho(\tau)\,u_0^2 \cos[\omega_0 t + \\ + \psi(t)]\cos[\omega_0(t-\tau) + \psi(t-\tau)], \tag{7}$$

$$R_{n,}(t,\ \tau) = N_0\delta(\tau)\,u_0^2 \cos[\omega_0 t + \psi(t)]\cos[\omega_0(t-\tau) + \psi(t-\tau)] = KN_0\delta(\tau). \tag{8}$$

We are not interested in the high-frequency components of the oscillations, for they do not pass through the rangefinder filter; hence, in expression (7) we shall reject the terms corresponding to these components and deal with the correlation function

$$R^0_{s,}(t,\ \tau) = KP_{s,}\rho(\tau)\cos\left[B(\tau)\sin\Omega_M\left(t - \frac{\tau+\tau_0}{2}\right)\right]. \tag{9}$$

In formulas (8) and (9) we introduce the designations

$$\left.\begin{aligned} K &= \frac{u_0^2}{2}, \\ P_s &= \frac{u^2}{2}, \\ B(\tau) &= 4\frac{\Delta\omega}{\Omega_M}\sin\Omega_M\frac{\tau_0}{2}\sin\Omega_M\frac{\tau}{2} \cong 2\Delta\omega\tau_0\sin\Omega_M\frac{\tau}{2}, \end{aligned}\right\} \tag{10}$$

where P_s is the average power of the received signal and K is a conversion factor in which we may include the gain of the following sections.

A random signal with correlation function (6) passes through the filter. We shall assume that this is an ideal filter with frequency characteristic of Figure 2. With $\Delta\omega\tau_0 >> 1$ the principal spectral components of a random signal with correlation function (9) are located in the region of the filter's passband and hence the "signal" component of the correlation function of the process at the filter output is represented with sufficient accuracy by

$$R_s(t,\ \tau) = R^0_{s,}(t,\ \tau). \tag{11}$$

The "noise" component of the correlation function is defined as

$$R_n(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} KN_0\,|H_\Phi(j\omega)|^2 e^{j\omega\tau}\,d\omega \tag{12}$$

and after calculation of the correlation function of the process at the filter output

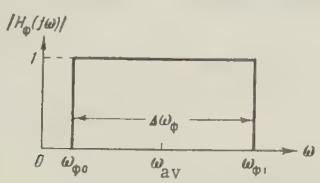


Figure 2. Frequency characteristic of filter.

$$R_{s,n}(t,\ \tau) = R_s(t,\ \tau) + R_n(\tau) = \\ = KP_s\rho(\tau)\cos\left[B(\tau)\sin\Omega_M\left(t - \frac{\tau+\tau_0}{2}\right)\right] + \\ + KP_n\left(\frac{\sin\omega_{\Phi 1}\tau}{\Delta\omega_\Phi\tau} - \frac{\sin\omega_{\Phi 0}\tau}{\Delta\omega_\Phi\tau}\right), \tag{13}$$

where

$$P_n = 2\Delta f_\Phi N_0 \tag{14}$$

represents the average noise power in the filter passband. Thus, we have found the correlation function of the random process at the filter output, the zero count of which defines the output value of the range finder.

2. SYSTEMATIC ERROR OF RANGEFINDER

For calculation of the systematic error due to noise and signal fluctuations it is necessary to find the average number of zeroes of the process at the rangefinder filter output. during $T(T \gg \frac{2\pi}{\Omega_M})$ on the condition of a positive derivative of the process at the zeroes.

The correlation function of the examined process is defined by expression (13) The variance of this process

$$\sigma^2 = R_{s,n}(t, 0) = K(P_s + P_n) \tag{15}$$

does not depend on current time t.

The correlation coefficient

$$r_{s,n}(t, \tau) = \frac{R_{s,n}(t, \tau)}{\sigma^2} = \frac{P_s}{P_s + P_n} \rho(\tau) \cos\left[B(\tau) \sin \Omega_M \left(t - \frac{\tau + \tau_0}{2}\right)\right] + \\ + \frac{P_n}{P_s + P_n} \left(\frac{\sin \omega_{\Phi 1}\tau}{\Delta\omega_\Phi \tau} - \frac{\sin \omega_{\Phi 0}\tau}{\Delta\omega_\Phi \tau}\right). \tag{16}$$

It is easily shown that the case of a nonstationary random process with a variance independent of time may be dealt with by the method for calculation of the average number of zeroes, proposed, for example, by B.R. Levin [2] for the case of a stationary random process. As a result the average number of zeroes per unit time in weighting the system over time T is defined as

$$n_{av} = \frac{1}{T} \int_t^{t+T} \frac{\omega_I(t)}{2\pi} dt \tag{17}$$

and $\omega_I(t)$ is defined by means of

$$\omega_I^2(t) = -r''(t, 0), \tag{18}$$

where $r(t, \tau)$ is the correlation coefficient of the examined process and

$$r''(t, 0) = \frac{d^2 r(t, \tau)}{d\tau^2}\Big|_{\tau=0} \tag{19}$$

In the case of a nonstationary process ω_I is a time function.

Calculating $\omega_I^2(t)$, we have

$$\omega_I^2(t) = -r''_{s,n}(t, 0) = \frac{P_s}{P_s + P_n}\left[-\rho''(0) + (\Delta\omega\Omega_M\tau_0)^2 \sin^2 \Omega_M\left(t - \frac{\tau_0}{2}\right)\right] + \\ + \frac{P_n}{P_s + P_n}\left(\omega_{av}^2 + \frac{\Delta\omega_\Phi^2}{12}\right), \tag{20}$$

where

$$\omega_{av} = \omega_{\Phi 0} + \frac{\Delta\omega_\Phi}{2} = \omega_{\Phi 1} - \frac{\Delta\omega_\Phi}{2}, \quad \rho''(0) = \frac{d^2\rho(\tau)}{d\tau^2}\Big|_{\tau=0}. \tag{21}$$

If we designate

$$\alpha = P_c\left[-\rho''(0) + \frac{(\Delta\omega\Omega_M\tau_0)^2}{2}\right] + P_n\left(\omega_{av}^2 + \frac{\Delta\omega_\Phi^2}{12}\right), \quad \beta = P_s\frac{(\Delta\omega\Omega_M\tau_0)^2}{2}, \tag{22}$$

then, using formula (17),

$$n_{av} = \frac{\sqrt{2}}{2\pi\sqrt{P_s + P_n}}\left(1 - \frac{1}{16}\frac{\beta^2}{\alpha^2} - \frac{5}{1024}\frac{\beta^4}{\alpha^4} - \cdots\right). \tag{23}$$

With $\frac{\beta}{\alpha} < 1$ the derived series shows excellent convergence. Hence, with sufficient accuracy

$$n_{av} = \frac{\sqrt{\alpha}}{2\pi\sqrt{P_s + P_n}} = \\ = \frac{1}{2\pi}\sqrt{\frac{P_c}{P_s + P_n}\left[-\rho''(0) + \frac{(\Delta\omega\Omega_M\tau_0)^2}{2}\right] + \frac{P_s}{P_n + P_n}\left(\omega_{av}^2 + \frac{\Delta\omega_\Phi^2}{12}\right)}. \tag{24}$$

In the absence of receiver noise $P_n = 0$, and in the absence of fluctuation $\rho''(0) = 0$. Hence in the case of a deterministic process at the receiver input

$$n_{av} = \frac{1}{2\pi} \frac{\Delta\omega\Omega_M}{\sqrt{2}} \tau_0, \tag{25}$$

and the output value of the rangefinder is, in effect proportional to the measured signal delay τ_0 with proportionality coefficient

$$k_\tau = \frac{1}{2\pi} \frac{\Delta\omega\Omega_M}{\sqrt{2}}. \tag{26}$$

The average value of delay measured in the presence of noise and fluctuation

$$\tau_{meas} = \frac{n_{av}}{k_\tau} = \sqrt{\frac{P_s}{P_s + P_n}\left[\frac{-2\rho''(0)}{(\Delta\omega\Omega_M)^2} + \tau_0^2\right] + \frac{P_n}{P_s + P_n} \frac{2}{(\Delta\omega\Omega_M)^2}\left(\omega_{\Phi av}^2 + \frac{\Delta\omega_\Phi^2}{12}\right)} \tag{27}$$

while the average value of measured range

$$D_{meas} = \frac{c\tau_{meas}}{2}, \tag{28}$$

where c is the velocity of light.

The relative systematic error in measurement of range is expressed by the formula

$$\frac{\Delta D}{D} = \frac{\Delta\tau_0}{\tau_0} = \\ = \sqrt{\frac{P_s}{P_s + P_n}\left[\frac{c^2\Delta\omega_{fl}^2}{2(\Delta\omega\Omega_M D)^2} + 1\right] + \frac{P_n}{P_s + P_n} \frac{c^2}{2(\Delta\omega\Omega_M D)^2}\left(\omega_{\Phi av}^2 + \frac{\Delta\omega_\Phi^2}{12}\right)} - 1 \tag{29}$$

and varies with the range.

In expression (29) $\Delta\omega_{fl}$ represents the quantity $\sqrt{-\rho''(0)}$, coinciding with the bandwidth (mean-square) of fluctuation in the reflected signal. With sufficient accuracy for practical purposes $\Delta\omega_{fl}$ may be taken to represent the fluctuation bandwidth determined by any method.

3. FLUCTUATION ERROR OF RANGEFINDER

In addition to the systematic error, noise and signal fluctuation cause a fluctuation error which may be characterized by the variance of the output value of the rangefinder. For exact calculation of this dispersion it would be necessary to extend to the case of a nonstationary random process the calculation of the correlation function and spectral density of the derivative from the phase of the random process. It may be shown that in the case of time-independent variance of the investigated process V.I. Bunimovich's formulas, (see reference [3]), derived for a stationary process, prove to be valid, and only the spectral density depends on time as a parameter by virtue of the nonstationary nature of the process.

It is necessary to point out, however, that the mentioned method of calculation requires specific treatment of the form of the fluctuation correlation function and leads to unwieldy results. Hence we shall limit calculation to an approximate evaluation of error variance. This evaluation is based on the following considerations.

The correlation coefficient of the process at the filter output (16) may be rewritten in the form

$$r_{s,n}(t, \tau) = \lambda(t, \tau)\cos\omega_{av}\tau + \chi(t, \tau)\sin\omega_{av}\tau, \tag{30}$$

where

$$\lambda(t, \tau) = \frac{P_s}{P_s + P_n}\rho(\tau)\cos[b(t, \tau)] + \\ + \frac{P_n}{P_s + P_n} \frac{\sin(\omega_{\Phi 1} - \omega_{av})\tau - \sin(\omega_{\Phi 0} - \omega_{av})\tau}{\Delta\omega_\Phi \tau}, \tag{31}$$

$$\chi(t, \tau) = -\frac{P_s}{P_s + P_n}\rho(\tau)\sin[b(t, \tau)] + \\ + \frac{P_n}{P_s + P_n} \frac{\cos(\omega_{\Phi 1} - \omega_{av})\tau - \cos(\omega_{\Phi 0} - \omega_{av})\tau}{\Delta\omega_\Phi \tau}, \tag{32}$$

and

$$b(t, \tau) = B(\tau) \sin \Omega_M \left(t - \frac{\tau + \tau_0}{2}\right) - \omega_{av}\tau. \tag{33}$$

In these expressions ω_{av} is a certain average frequency of the random process at the filter output.

Extending to the case of a nonstationary normal random process with time-independent variance the method of calculation presented in reference [4] for a unidimensional law of distribution for the derivative of the phase of the process, we obtain

$$W(\dot{\Theta}) = \frac{\delta\omega^2(t)}{2\{[\omega_{av} - \omega_I(t) - \dot{\Theta}]^2 + \delta\omega^2(t)\}^{3/2}}, \tag{34}$$

where

$$\delta\omega^2(t) = \omega_I^2(t) - \omega_{II}^2(t), \tag{35}$$

$$\omega_I^2(t) = -r''_{s,n}(0), \tag{36}$$

$$\omega_{II}(t) = \lambda(t, 0)\omega_{av} - \chi'(t, 0). \tag{37}$$

Inserting into formula (35) the expressions derived above, we have

$$\delta\omega^2(t) = \frac{P_s}{P_s + P_n}|\rho''(0)| + \frac{P_s P_n}{(P_s + P_n)^2}(\Delta\omega\Omega_M\tau_0)^2 \sin^2 \Omega_M\left(t - \frac{\tau_0}{2}\right) - \\ - \frac{2P_s P_n}{(P_s + P_n)^2}\Delta\omega\Omega_M\tau_0\omega_{av} \sin \Omega_M\left(t - \frac{\tau_0}{2}\right) + \\ + \frac{P_n P_c}{(P_s + P_n)^2}\omega_{av}^2 + \frac{P_n}{P_s + P_n}\frac{\Delta\omega_\Phi^2}{12}. \tag{38}$$

In obtaining the variance corresponding to the probability density (34) and averaging it over time (by virtue of the averaging properties of the subsequent elements of the circuit), we would obtain data for calculation of the variance of the output value of the rangefinder. However, the variance corresponding to expression (34) reverts to ∞. Hence, the width of the probability density curve is conveniently evaluated, according to B.R. Levin [2], by the quantity

$$\int_{-\infty}^{\infty} |x| \frac{\delta\omega^2(t)}{2[x^2 + \delta\omega^2(t)]^{3/2}} dx = \delta\omega(t). \tag{39}$$

We may see that the spectral density of the process of the zero count in the low-frequency region is approximately

$$S_0 = k\frac{\overline{\delta\omega(t)}}{2\pi}, \tag{40}$$

where the superior bar denotes time averaging and k is a coefficient of the order of unity. Assuming k = 1 and making the approximate substitution

$$\sigma_{out}^2 \cong \frac{1}{2\pi T}\sqrt{\frac{P_s}{P_s + P_n}|\rho''(0)| + \frac{P_s P_n}{(P_s + P_n)^2}\left[\frac{(\Delta\omega\Omega_M\tau_0)^2}{2} + \omega_{av}^2\right] + \frac{P_n}{P_s + P_n}\frac{\Delta\omega_\Phi^2}{12}} \tag{41}$$

and the variance in range measurement may be approximately evaluated from the formula

$$\sigma_D^2 \cong \frac{\pi c^2}{T(\Delta\omega\Omega_M)^2}\sqrt{\frac{P_s}{P_s + P_n}\Delta\omega_{fl}^2 + \frac{P_s P_n}{(P_s + P_n)^2}\left[2\left(\frac{\Delta\omega\Omega_M D}{c}\right)^2 + \omega_{av}^2\right] \cdots \rightarrow} \\ \cdots \rightarrow \sqrt{+\frac{P_n}{P_s + P_n}\frac{\Delta\omega_\Phi^2}{12}}. \tag{42}$$

4. EVALUATION OF RESULTS

The formulas derived for systematic and fluctuation error are conveniently presented in the form

$$\frac{\Delta D}{D} = \sqrt{\frac{\varkappa \frac{\Delta\omega_{f1}^2}{d^2} + 1 + 0.27\varepsilon d^2}{1 + \varepsilon d^4}} - 1 \tag{43}$$

and

$$\sigma_D^2 = \frac{2\pi}{T} \varkappa D_m^2 \sqrt{\frac{\Delta\omega_{f1}^2}{1 + \varepsilon d^4} + \frac{\varepsilon d^4}{(1 + \varepsilon d^4)^2} \frac{d^2 + \frac{2}{\pi^2}}{\varkappa} \frac{\varepsilon d^4}{1 + \varepsilon d^4} \frac{2}{3\pi^2} \frac{1}{\varkappa}}, \tag{44}$$

where the introduced designations are: D_m — the maximum effective range;

$$d = \frac{D}{D_m}; \quad \varkappa = \frac{\nu^2}{2\pi^2 \Delta\omega^2}; \quad \nu = \frac{2\pi}{\Omega_M \tau_{0m}}$$

— the ratio of the modulation period to the maximum delay of the signal, and

$$\varepsilon = \frac{P_n}{P_s}$$

in the ratio of the modulation period to the maximum range.

In deriving expressions (43) and (44) it was assumed that $\omega_{f1} >> \omega_{f0}$ and hence $\Delta\omega_f \approx \approx \omega_f \approx 2\omega_{fav}$. In addition, it was considered that the minimum value of $\Delta\omega_f$ must be determined from passband conditions corresponding to maximum range, hence,

$$\Delta\omega_\Phi \cong \frac{4\Delta\omega}{\nu}.$$

From formula (43) it is seen that at short ranges (d<<1) the average value of measured range will be greater than actual range. However, with the choice of sufficiently high frequency deviation $\Delta\omega$ and modulation frequency Ω_M, the inequality

$$\varkappa \frac{\Delta\omega_{f1}^2}{d^2} \ll 1 \tag{45}$$

applies even at short range. Hence, the systematic error at short range is not great. At long range the average value of measured range will be less than the actual value and at maximum range may be approximately defined by the formula

$$\frac{\Delta D}{D_m} \cong \sqrt{\frac{1 + 0.27\varepsilon}{1 + \varepsilon}} - 1.$$

If it is assumed that at maximum range the noise is considerably greater than the signal ($\epsilon >> 1$), then from formula (46) $\frac{\Delta D}{Dm} = 0.48$, which is inadmissable. With $\epsilon = \frac{\Delta D}{Dn} = -0.2$; that is, the measured range is 20% smaller than the actual range. Assigning a permissible error of $\frac{\Delta D}{D_n} = -\mu$ in measuring the average falue of range, we have

$$\varepsilon = \frac{1 - (1 - \mu)^2}{(1 - \mu)^2 - 0.27}. \tag{47}$$

If, for example, we assign $\mu = 0.05$, then $\epsilon = 0.159$. Thus, we find the permissible signal-to noise value from which we may determine the required transmitter power and receiver sensitivity.

From expression (45) it is seen that with $\epsilon < 1$ the fluctuation error is maximum at maximum range D_m and may be calculated from the formula

$$\sigma_{D\,max}^2 = \frac{2\pi}{T} \varkappa D_m^2 \sqrt{\frac{\Delta\omega_{f1}^2}{1 + \varepsilon} + \frac{1}{\varkappa} \frac{\varepsilon}{1 + \varepsilon} \left(\frac{1,2}{1 + \varepsilon} + 6.75 \cdot 10^{-2}\right)}. \tag{48}$$

Upon observing the often satisfied conditions $\Delta\omega_{fl}^2 \varkappa \ll 1$ and $\epsilon < 1$ the following approximate formula is applicable

(49)

$$\sigma_{D\max}^2 \cong \frac{\sqrt{2.4}\nu D_m^2}{T\Delta\omega}\frac{\sqrt{\varepsilon}}{1+\varepsilon} \qquad \sigma_0^2 = \frac{\sigma_{D\max}^2}{\frac{\sqrt{2.4}\nu D_m^2}{T\Delta\omega}} = \frac{\sqrt{\varepsilon}}{1+\varepsilon},$$

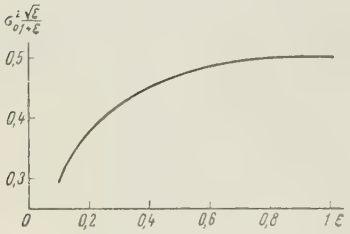


Figure 3. Maximum variance of rangefinder error as a function of signal-to-noise ratio at maximum range ϵ.

Figure 3 is the plot of a curve from which we may determine the permissible values of signal-to-noise ratio for a given variance of range-finder output.

It is easily seen that in order to decrease the fluctuation error it is desirable to increase the frequency deviation $\Delta\omega$, averaging time T, modulation frequency Ω_M (to decrease ν) and also, it appears, to increase the signal-to-noise ratio (to decrease ϵ).

If, for example, we choose $\Delta f = 50$ Mc, $T = 0.2$ sec, and $\nu = 10$, then for $D_m = 10$ cm.

$$\sigma_{D\max}^2 = 13{,}1\,\text{m}^2(\varepsilon = 1),$$

$$\sigma_{D\max}^2 = 8{,}7\,\text{m}^2(\varepsilon = 0{,}159).$$

In this case it is not difficult to see that for Δf_{fl} of the order of tens of cps the conditions of validity of the last formula are satisfied.

Thus, with the usual parameter relationships the errors in such rangefinders are small, It must also be pointed out that the examined rangefinder circuit is far from optimum. It may be brought to near-optimum by making an extremely narrowband filter tunable according to the measured range. It is evident that this is effective only in high-precision systems or in systems operating at small signal-to-noise ratios.

The author wishes to express his thanks to V.G. Repin for a number of valuable comments on the work presented here.

REFERENCES

1. B.V. Malanov, Toward a theory for a radio rangefinder with frequency modulation, Radiotekhnika, 1956, 14, 4, 46.
2. B.R. Levin, Teoriya sluchaynnykh protsessov i yeye primeneniye v radiotekhnike, [Theory of random processes and its application in radio engineering]; Izd, Sivetskoye radio, 1957.
3. V.I. Bunimovich, Flyuktuatsionnyye protsessy v radiopriyemnikh ustroystvakh, [Fluctuation processes in radio receivers], Izd. Sovetskoye Radio, 1951.

Submitted to the editors 28 January 1960

LATERAL RADIATION OF PARABOLIC ANTENNAS

B. Ye. Kinber

The article discusses a method of calculating lateral radiation of mirror antennas which is based on asymptotic solutions of diffraction problems. It is shown that previous works on the theory of mirror antennas have not considered a number of important phenomena and, hence, in some cases have presented incorrect solutions. Curves are given for lateral-radiation envelopes of cylindrical and axially symmetric antennas.

INTRODUCTION

Radiation calculation of mirror antennas is presently performed by aperture and current methods. For these calculations, it is necessary to know either the distribution of the field at the aperture or the distribution of current over the mirror. Disregarding difficulties in calculation, accuracy in calculating the radiation field is determined by the accuracy of definition of the initial distributions

However, distribution of current over the surface and distribution of the field at the aperture are usually given in the Kirchhoff approximation (sometimes with a few refinements) and are not obtained from strict solutions; hence, they limit the accuracy of calculation.

Particularly unfavorable is the fact that the actual accuracy of the initial approximation and the boundaries of its applicability are unknown; this is also true of the corrections required for calculation in the higher approximation.

In addition to this, the complex mathematical apparatus used in the aperture and current methods creates a false impression of exactness, in which connection the results of calculation are often extended to a region of parameters where the initial approximations are unacceptable.

The present report discusses a method of calculation of lateral radiation of a mirror based on the use of asymptotic solutions of diffraction problems.

Lateral radiation of a mirror antenna is defined in the form of a sum of "rays" satisfying the Fermat principle, being a well-known generalization of the Keller method, which introduced the concept of diffraction rays bounding off the edges [1] and enveloping the body.

In the case of a mirror of finite dimensions, along with the Keller rays, there are more complex types of rays corresponding to combinations of multiple reflections over the concave portion of the mirror, diffractions at the edges, bends around the rear of the mirror, etc.

Each ray corresponds to a field propagating along the ray; amplitude of the field is a slowly varying function of points of radiation and the observation points and shape of the body. Due to interference of the rays the resulting field has an oscillating character.

Section 1 defines the possible types of rays, then discusses the amplitudes of the fields corresponding to the various types of rays. In radiation pattern calculation with an accuracy of $(kD)^{-3/2}$ it suffices to consider only the rays whose amplitude depends on wavelength not greater than $\lambda^{1/2}$. This limits the number of rays which need to be considered in calculation. With this degree of accuracy the envelope of the characteristic of lateral radiation of mirror antennas is calculated.

It is understood that in the present report the stated problem is far from exhausted and that further refinements are required. However, the solution as derived is more accurate than that obtained by the current and aperture methods.

Among the advantages of this solution are the physical descriptiveness and simplicity of calculations.

In particular, it follows from the present work that the influence of the edge of the

mirror (sharp rim) on the distribution of currents along its surface, in distinction from the usual representations, is not localized near the edge, but extends over the entire surface of the mirror. The value of correction for current $2[\vec{n}\,\vec{H}_0]$ is commensurate with and in certain cases exceeds this current.

1. DIFFRACTION RAYS AND THEIR AMPLITUDES

Let us examine all the possible types of extreme paths between source O and observation point P.

In addition to the direct path between them (O-P) there are possible extreme paths with the additional conditions: an extreme path must come into contact once or several times with a diffracting body (points R), sharp edges or points of the body (points H), or bend around the body (in arc s-s) (see Figures 1 and 2). We will consider only those extreme paths which do not pass through the body (points T in Figures 1 and 2). In the general case the extreme path consists of a sequence of several reflections R, contacts with the edges H and curves of the body s-s.

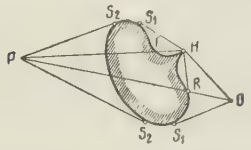


Figure 1. Extreme paths between source O and observation point P.

Let us characterize each extreme path as a sequence of points R, H, s. Thus, for example, a ray reflected from the mirror is characterized by the sequency O-H-P; a ray corresponding to diffraction at one edge is characterized by the sequence O-H-P. The extremum of the ray length is ensured by the fact that: (a) at points R the law of specular reflection applies; (b) at points H the angle between the border and the incident ray is equal to the angle between the border and the diffracted ray; (c) the tip of the body is always a point of an extreme path; (d) curvature s-s of the body occurs along geodetic lines. The incident and deflected rays are tangent to the body and arc s-s.

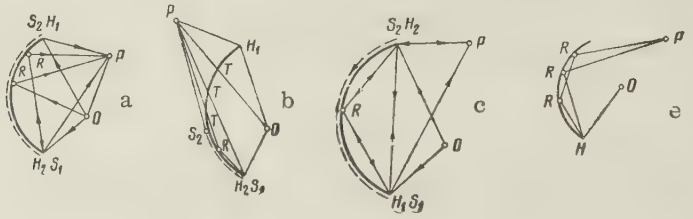


Figure 2. Extreme paths in diffraction at mirror.

a — $O-P$, $O-R-P$, $O-H_1-P$, $O-H_2-P$, $O-H_2S_1-S_2H_1-P$, $O-H_2-R-P$; b — $O-P$, $O-T-P$, $O-H_1-P$, $O-H_2S_1-S_2-P$, $O-H_2-T-P$; $O-H_2-R-T-P$; c — $O-H_2-H_1S_1-H_2S_2-H_1-P$, $O-H_1-H_2S_2-H_1S_1-P$, $O-H_1S_2-H_2S_2-P$, $O-H_2-H_1-P$, $O-H_1-H_2-P$, $O-H_1-R-H_2-P$ etc., d — $O-H-R-P$, $O-H-R-R-P$

The number of extreme paths depends on the shape of the body, the locations of the source and observation point. For bodies of infinite dimensions the number of extreme paths may be finite.

For example, if the source is located at the focus of the paraboloid, the extreme paths are always two (O-P and O-R-P). For bodies of finite dimensions the number of extreme paths is always infinite. In fact, to any path bending around the body there may always infinite. In fact, to any path bending around the body there may always be added an extreme path consisting of any number of revolutions around the body.

Let us now discuss the amplitudes of the various types of rays. A change in amplitude along a ray occurs between points of "contact with the body and at points (lines)of contact with the body. In the first approximation, with the exception of special cases, the change in amplitude along sections of rays between contacts with the body are subject to laws of geometric optics. Since the energy flux in the ray tube is preserved, the amplitude is proportional to $s^{-1/2}$, where s is the cross-sectional area of the tube and is independent of wavelength.

Table 1

Sector No.	Sector Boundary	Types of extreme paths	Note
1	$(O, (kD)^{-1})$	$O-P$ $O-R-P$ $O-H_1-H_2-P$ $O-H_2-H_1-P$ etc.	main lobe
2	$\left((kD)^{-1}, \frac{\pi}{2}-\psi_{cr}\right)$	$O-P$ $O-H_1-P$ $O-H_2-P$ $O-H_1-H_2-P$ $O-H_2-H_1-P$ etc.	sector of usual side lobes, interference of fringe waves and radiator field
3	$\left(\frac{\pi}{2}-\psi_{cr}, \frac{\pi}{2}-\frac{\psi_{cr}}{2}\right)$	$O-P$ $O-H_1-P$ $O-H_2-P$ $O-H_2-R-P$ $O-H_2-R-R-P$ etc.	side lobes formed due to interference of fringe waves and waves reflected from mirror surface
4	$\left(\frac{\pi}{2}-\frac{\psi_{cr}}{2}, \frac{\pi}{2}\right)$	$O-P$ $O-H_1-P$ $O-H_2-P$ $O-H_2-R-P$ etc.	part of reflected fringe rays shadowed
5	$\left(\frac{\pi}{2}, \pi-\psi_{cr}\right)$ or $\left(\frac{\pi}{2}, \frac{\pi}{2}+\frac{\psi_{cr}}{2}\right)$	$O-P$ $O-H_1-P$	wave from far edge of mirror shadowed
6	$\left(\pi-\psi_{cr}, \frac{\pi}{2}+\frac{\psi_{cr}}{2}\right)$ or $\left(\frac{\pi}{2}+\frac{\psi_{кр}}{2}, \pi-\psi_{cr}\right)$	$O-H_1-P$ or $O-P$ $O-H_1-P$ $O-H_2-P$	primary radiator and far edge of mirror shadowed
7a	$(\pi-\psi_{cr} \mp (kD)^{-1})$	--	sector of fringe lobe of radiator
7b	$\left(\frac{\pi}{2} \mp (kD)^{-1}\right)$		sector of fringe lobe of far edge of mirror
8	$\left(\frac{\pi}{2}-\psi_{cr}, \pi-(kD)^{-1}\right)$ or $\left(\frac{\pi}{2}+\frac{\psi_{cr}}{2}, \pi-(kD)^{-1}\right)$	$O-H_1-P$ $O-H_2-P$	shadowed radiator
9	$(\pi-(kD)^{-1}, \pi)$	$O-H_1-P$ $O-H_2-P$	fringe wave cophasal; sector of rear lobe

Table 2

Sector No.	$2\psi_{cr}^{\circ}$				
	60	90	120	150	180
2	0—60	0—45	0—30	0—15	0
3	60—75	45—67,5	30—60	15—52,5	0—45
4	75—90	67,5—90	60—90	52,5—90	45—90
5	90—105	90—112,5	90—120	90—105	90
7	105—150*	112,5—135*	--	105—127,5**	90—135*
8	150—180	135—180	120—180	127,5—135	135—180

* No shadowing
** Radiator and far edge shadowed

Note: Comma designates decimal point.

The change in amplitude at points of contact with the body occurs (in the first approximation) as follows.

(a) At points of reflection the solid angle of the ray tube changes. The expansion or contraction coefficient is determined by the geometry of reflection and is independent of wavelength.

(b) At points H a ray tube of width of the order of λ in the direction crosswise to the edge (width of the last Fresnel zone) is changed to a cylindrical or toroidal wave. At sharp points a ray tube with area of the order of λ^2 is transformed to a spherical wave. Consequently, at points H on the edges the transformation coefficient is proportional in amplitude to $\lambda^{1/2}$ and at sharp points is proportional in amplitude to λ.

(c) Along the bend line s-s of the body (along the convex side of the body) there occurs an exponential damping of amplitude. The exponential is proportional to the arc s-s.

These considerations suffice to permit selection of the extreme paths which contribute with an assigned amount of smallness.

The usual calculation of lateral radiation in the best case corresponds to consideration of single diffractions at the edges (O-H-P), that is, the consideration of terms of the order of $\lambda^{1/2}$ (or $kD^{-3/2}$) for a normalized pattern). If such accuracy is preserved but all the diffraction phenomena are more systematically considered, then it suffices to consider all the paths containing single diffractions at the edges.

Each subsequent diffraction at an edge reduces the order of amplitude by $\lambda^{1/2}$. We may also disregard rays bending around the mirror, particularly where the rear of the mirror is occupied by structural elements securing the antenna.

Thus, with the chosen accuracy of calculation $[(-kD)^{-3/2}]$ it suffices to consider rays of the type:)-P (field of primary radiator); O-H-P (diffraction field at rim of antenna); O-H-R-P; O-H-R-...-R-P (reflection of fringe waves from concave side of mirror).

In distinction from the current method, the calculation automatically allows for the shadowing condition (reference[3]), since only those extreme paths which do not intersect the mirror itself are chosen.

Table 1 presents a breakdown of the pattern of an axially symmetric parabolic mirror into sectors according to types of extreme paths and Table 2 lists in degrees the angles defining these sectors. The angle of radiation of the mirror (see Figure 6) is $2\psi_{cr}$.

The breakdown into sectors is based solely on geometric considerations. The sector boundaries correspond to transitional regions, chief of those listed in Table 1 being the main lobe of the radiation pattern, the rear lobe, the fringe lobes of the radiator and the far edge.

2. FRINGE WAVE (O-H-P)

The asymptotic of exact solutions of diffraction problems for bodies with a sharp edge (se, for example, reference [4]) beyond the light-shadow boundary contains a wave issuing from the edge -- cylindrical for a half-plane, toroidal for a disc. The presence of the fringe wave is not associated with the exactness of solution. A term corresponding to the fringe wave is also used in solution in the Kirchhoff approximation [5, 6].

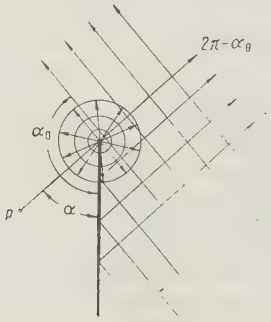


Figure 3. Analysis of fringe wave.

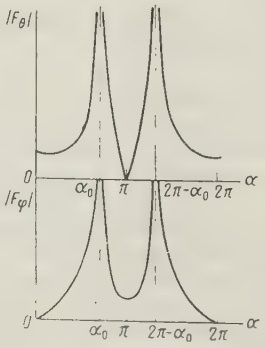


Figure 4. Curves of F_θ and F_φ.

The fringe wave is associated with radiation of currents flowing over the last Fresnel fringe zone at the mirror surface. Since the width of the zone depends on the observation point, the fringe wave has a "pattern" and depends on the details of current distribution in the Fresnel zone.

From the solution for the half-plane (reference [4]) it follows that for $kr \gg 1$ (i.e., at a distance of one or two wavelengths from the edge for a primary wave with vector H_0 parallel to the edge

$$\vec{H}_{cr} = \vec{H}_0 \sqrt{\frac{2}{\pi kr}} e^{i\left(kr + \frac{\pi}{4}\right)} F_\varphi;$$

$$F_\varphi = \frac{\sin\frac{\alpha_0}{2}\cos\frac{\alpha}{2}}{\cos\alpha - \cos\alpha_0}, \tag{1}$$

and for the incidence of a wave with the vector E_0 parallel to the edge

$$\vec{E}_{cr} = \vec{E}_0 \sqrt{\frac{2}{\pi kr}} e^{i\left(kr + \frac{\pi}{4}\right)} F_\varphi; \quad F_\varphi = \frac{\cos\frac{\alpha_0}{2}\sin\frac{\alpha}{2}}{\cos\alpha - \cos\alpha_0}, \tag{2}$$

where $\vec{H}_0$, $\vec{E}_0$ is the strength of the primary field at the fringe of the mirror, r is the distance from the edge to the observation point, α_0 is the angle of incidence and α is the angle of observation (see Figure 3).

Curves for F_θ and F_φ are given in Figure 4. Direction $2\pi - \alpha_0$ corresponds to the specularly reflected wave and α_0 corresponds to the incident wave. Let us write for comparison the expression for the fringe wave using the Kirchhoff approximation for a primary wave with the vector E_0 parallel to the edge

$$\vec{E}_{cr} = \vec{E}_0 \sqrt{\frac{2}{\pi kr}} e^{i\left(kr + \frac{\pi}{4}\right)} F_K; \quad F_K = \frac{1}{2}\frac{\sin\alpha_0}{\cos\alpha - \cos\alpha_0}. \tag{3}$$

It is of interest to note that if the angles are read from the direction of the reflected wave

$$\alpha = 2\pi - \alpha_0 - \theta, \tag{4}$$

then in the Kirchhoff approximation

$$F_K = \frac{1}{2}\frac{\sin\alpha_0}{\sin\theta\sin\alpha_0 + \cos\alpha_0(1 - \cos\theta)},$$

and for $\theta \ll 1$ we obtain the pattern of the fringe wave in the aperture approximation:

$$F_a = \frac{1}{2\sin\theta}. \tag{5}$$

With $\theta \ll 1$ formula (5) also follows from formulas (1) and (2).

The fringewave field, generated by currents at the mirror does not convey the character of radiation of these currents in the direction of the main lobe and fringe lobe where F_φ and F_θ assume infinitely large values. In these directions, instead of using the fringe waves, it is convenient to represent the total field in the form of an integral for the total system of currents over the mirror (both for the uniform and for the nonuniform components: see, for example, reference [6]).

Since the lateral radiation (without considering multiple reflections, etc) is due to interference of fringe waves originating at opposite edges of the mirror, then from (1)-(5) it follows that: (a) the envelope of the lobes of lateral radiation does not decrease monotonically with an increase in θ, since $F(\theta)$ are not monotonic functions of θ; (b) the envelope of lateral radiation is determined by only orientation of the edges of the mirror (i.e., it does not depend on the general shape of the mirror).

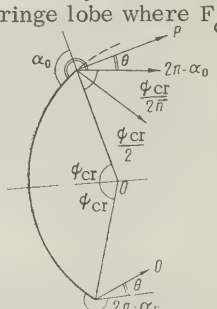


Figure 5. Relationship between angles α_0, ψ_{cr}, and θ in parabolic mirror.

Consequently, with various mirror shapes but with identical orientation of the edges, the lateral radiation remains the same; (d) the mirror does not possess a phase center.

Let us explain the last remark. Let us express α and α_0 in terms of ψ_{cr} and θ (Figure 5).

$$\alpha_0 = \frac{\pi}{2} + \frac{\psi_{cr}}{2}, \tag{6}$$

$$\alpha = 2\pi - \alpha_0 - \theta \tag{7}$$

for the upper edge of the mirror and

$$\alpha = 2\pi - \alpha_0 + \theta \tag{8}$$

for the lower edge.

For $\theta \ll 1$ the minimum of the lobes is other than zero

$$f_{E\min} \simeq \frac{1}{2} \operatorname{tg} \frac{\alpha_0}{2} = \frac{1}{2}\left(\frac{\pi}{4} + \frac{\psi_{cr}}{2}\right) \neq 0; \quad f_{H\min} \simeq \frac{1}{2} \operatorname{ctg}\left(\frac{\pi}{4} + \frac{\psi_{cr}}{2}\right) \neq 0 \tag{9}$$

for the E-plane and the H-plane.

Consequently, the pattern of the mirror has no zeroes.

The presented calculations have pertained to a cylindrical mirror in which the fringe wave is purely cylindrical. For an axially symmetrical mirror with a point radiator the cylindrical fringe wave must be replaced by a toroidal wave which, in the final analysis goes over into a spherical wave. Using the stationary-phase method, it is easily shown that in calculating the pattern of the mirror the pattern of the fringe wave for the half-plane at the stationary point of the contour must be multiplied by $(\sin\theta)^{-1/2}$. It was assumed above that the rim of the mirror has a sharp edge and that illumination of the rim of the mirror is finite. For "zero" illuminations of the rim the intensity of the fringe wave and its pattern are changed.

Let us now evaluate the influence of curvature of the rim of the mirror. If $k\rho > 2$-3 (where ρ is the radius of curvature of the rim) then calculation of the fringe wave may be performed from the laws of geometric optics. From these it follows that the amplitude of the fringe wave for a mirror with a rounded rim does not depend on λ, that is, it is greater than at a mirror with a sharp rim and that the corresponding rays diverge in the direction away from the mirror axis. Hence, rays from different ends do not interfere with one another and the lateral radiation has a monotonic character. In the shadowing region ($\theta > \pi - \psi_{cr}$) the field of the fringe wave is smaller than in the case of a mirror with a sharp rim. In mirrors with rounded rims multiple reflections are not present.

3. MULTIPLE REFLECTIONS OF THE FRINGE WAVE (O-H_2-R-P, O-H_2-R-R-P, etc)

At a distance of $(1\text{-}2)\lambda$ from the edge the fringe wave may be regarded as a wave radiated by a current filament located at the edge. Analysis of such a source for the cases of a cylinder and a sphere shows that asymptotically the field in the vicinity of the concave surf

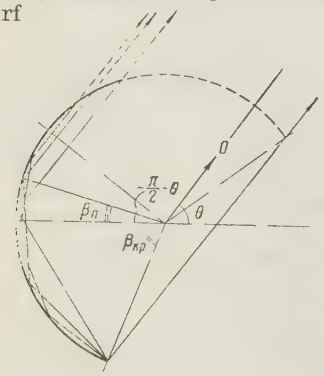


Figure 6. Rays reflected at angle θ.

surface may be regarded as the sum of rays of geometric optics and the field of the surface wave may be regarded as pressed to the concave wall ("whispering-gallery" effect). Since we are not discussing short diffractions, we may, with the exception of a small sector of angles with $\theta \simeq \frac{\pi}{2} - \frac{\psi_{cr}}{2}$, limit analysis to reflection by geometric optics.

Let us first discuss multiple reflections of the fringe wave in a spherical reflector, since it is close in shape to a parabolic mirror and the geometry of multiple reflections in it is simpler. Let us investigate the influence of multiple reflections on the far field (that is, on the pattern).

Figure 6 shows the structure of all rays departing from the mirror at angle θ. The single-reflection ray (O-H-R-is not shadowed. Rays of double and triple reflection

(O-H-R-R-T-P, O-H-R-R-R-T-P) are shadowed.

These rays we obtain by extending the n+1th side of an open rectilineal polygon based on an angle of $2(\beta_{cr} + \pi/2 - \theta)$ and consisting of 2n+1 members. Angle of the point of reflection of the ray is defined by the relationship

$$\beta_n = \left(\beta_{cr} + \frac{\pi}{2} - \theta\right)\left(1 - \frac{1}{2n+1}\right) - \beta_{cr}. \tag{10}$$

Boundary $\beta_n = \beta_{cr}$ defines the sector with n reflected rays. It is seen from (10) that if sector $\frac{\pi}{2} - \psi_{cr} < \theta < \frac{\pi}{2} - \frac{\psi_{cr}}{2}$ of multiple reflections is assumed as the unit, then segment (0, 1/2) is occupied by triple reflections, etc.

The sector of multiple reflections with shadowing $(\frac{\pi}{2} - \frac{\psi_{cr}}{2}, \frac{\pi}{2})$ is divided somewhat differently. If we reckon from angle $\pi/2$, then segment (0, 1/2) generally does not contain multiple reflections; (1/2, 2/3) are single, (1/3, 3/4) are double reflections, etc. (see Figure 7). Reflections of greater multiples in the first case are absent and in the second case are shadowed by the mirror itself.

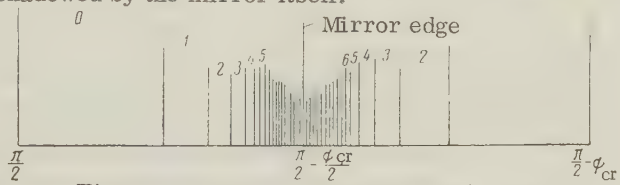


Figure 7. Sectors of multiple reflection.

Angle χ of ray departure from point H_2 to infinity at angle θ, after n reflections is defined by the formula

$$\chi = 2\pi - \frac{\beta_{cr} + \frac{\pi}{2} - \theta}{2n+1}, \tag{11}$$

which determines the widening of the ray after n reflections:

$$\nu = \frac{d\theta}{d\chi} = 2n + 1. \tag{12}$$

Thus amplitude of the n-th reflected ray decreases by $(2n + 1)^{-1/2}$ times. The path traveled by the ray (relative to the path from the center of the circle A) is

$$L \simeq R_0\left(\beta_{cr} + \frac{\pi}{2} - \theta\right)\left[1 - \frac{1}{6}\left(\frac{\beta_{cr} + \frac{\pi}{2} - \theta}{2n+1}\right)^2\right].$$

Cofactor $R_0\left(\beta_{cr} + \frac{\pi}{2} - \theta\right)$ is equal to the path along the arc; the second cofactor shows how much shorter the path along the chord is than the path along the arc. Actually the difference between the path along the chord and the path along the arc for usual wavelengths is small for rays which are reflected twice or more, and, hence, all reflected rays may be considered together. However, in adding rays of different multiples it is also necessary to consider that each reflection results in an additional phase π, and passage of a ray through the caustic results in a phase jump of $\pi/2$. Consequently, the total amplitude of all reflected rays is proportional to the field from a single-reflected ray and in rough calculation of the pattern with the desired degree of accuracy it is sufficient to consider only the single-reflected ray, assuming

$$\left|\sum_{q=1}^{n} \frac{e^{-i\frac{\pi}{2}q}}{\sqrt{2q+1}}\right| \simeq \frac{1}{\sqrt{3}}.$$

In the case of a cylindrical mirror the single-reflected ray is represented by the field:

$$\frac{1}{\sqrt{3}}\sqrt{\frac{2}{\pi kr}}\,e^{i\left(kl-\frac{\pi}{4}\right)}F\left(2\pi-\frac{\beta_{cr}+\frac{\pi}{2}-\theta}{3}\right), \tag{13}$$

proportional to the field of the fringe wave in the direction of ray departure before reflection and inversely proportional to the root of the ray expansion coefficient after reflection. For an axially symmetric mirror with a needle-shaped ray it is more convenient to examine the cone of the rays departing from the edge in the interval of angles (χ, $\chi + d\chi$) and the cone of the rays after reflection departing in the interval of angles (θ, $\theta + d\theta$). Representing the energy-balance equation for these pencils in the form

$$\frac{F^2(\chi)}{\cos(\beta_{cr}-\chi)}\cos(\beta_{cr}-\chi)\,2\pi\,d\chi = F^2(\theta)\sin\theta\,d\theta.$$

we obtain

$$F(\theta)=\frac{F(\chi)}{\sqrt{\chi\sin\theta}}. \tag{14}$$

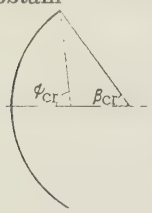


Figure 8. Illustrating the relationship between spherical and parabolic mirrors.

From this, it follows that, as in the case of cylindrical mirrors, after reflection the amplitude decreases by $\nu^{-1/2}$ times.

In calculating reflected rays in a parabolic mirror we may use formulas (13) and (14), in which is necessary only to express β_{cr} in terms of ψ_{cr}. Since a spherical mirror is almost parabolic, with the source located at the midpoint of the sphere radius (see Figure 8) β_{cr} is determined from the equation

$$\sin\psi_{cr}=\frac{\sin\beta_{cr}}{\sqrt{\frac{5}{4}-\cos\beta_{cr}}}. \tag{15}$$

The precise relationships for multiple reflections in a paraboloid (see Figure 9) may be presented in the form of recurrence relationships

$$\chi_n+\chi_{n+1}=\frac{\beta_{n+1}}{2}-\frac{\beta_n}{2}, \tag{16}$$

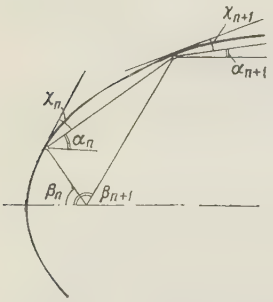


Figure 9. Illustrating the geometry of multiple reflections in a paraboloid of rotation.

For glancing reflections $\chi \ll 1$ the glance angle is almost unchanged

Dependence of α_{n+1} on χ_n is defined by the formula

4. GENERAL CHARACTERISTICS OF LATERAL RADIATION

From the foregoing analysis it follows that with an accuracy of up to $(kD)^{-3/2}$ lateral radiation of the mirror is formed due to four components: the fringe waves issuing from the far and near edges of the mirror, the fringe wave single-reflected from the mirror surface, and the radiator field.

These components have different phases and, if more than two of them have commensura

absolute values, the resulting field does not possess a clearly periodic lobe structure. In regions where the lateral radiation is formed chiefly due to one ray it is almost monotonic. The number of rays depends on the shadowing effect of the mirror.

For practical purposes greatest interest lies in the upper and lower envelopes of the interference structure of radiation rather than in the interference structure itself.

Obtaining the sum and difference of the absolute values, let us represent the expression for the envelopes of the normalized pattern of an axial-symmetric mirror with finite illumination of the edge, neglecting the influence of the radiator field (its lateral radiation) in the form

$$F(\varphi, \theta) = \sqrt{\frac{2}{\pi}} (kD)^{-1/2} (\sin\theta)^{-1/2} \left| F(-\theta) \mp \tilde{\varepsilon}_1 F(\theta) \mp \right.$$

$$\left. \mp \tilde{\varepsilon}_2 \frac{1}{\sqrt{3}} F(\chi) \right| \frac{T(\psi_{cr})}{r(\psi_{cr})} \frac{r(0)}{T(0)} \tau(\varphi), \qquad (21)$$

where $\frac{T(\psi_{cr})}{r(\psi_{cr})} \frac{r(0)}{T(0)}$ is the relative illumination of the edge of the mirror; $T(\varphi)$ is the excitation coefficient of a fringe wave of given polarization for points of the contour in the plane φ[1]; ϵ_1; and ϵ_2 are functions assuming values of 0 or 1 and considering the presence of absence of a ray; $F(-\theta)$. $F(t)$ is the pattern of a fringe wave of given polarization for the near and far end as a function of direction; $F(\chi)$ is the pattern of the fringe wave for a ray with reflection.

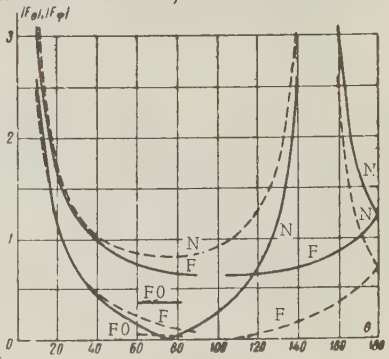


Figure 10

Let us explain the individual terms of expression (21). The value of relative illumination must not be less than 0.3 - 0.2, for in this case the dependence of f on kD and the form of functions of F are changed.

For the φ-component of polarization of lateral radiation F must be taken to be F_φ, defined by formula (2); wherein [see (6), (7), (8), (11)] $\alpha_0 = \frac{\pi}{2} + \frac{\psi_{cr}}{2}$, for $F(-\theta)$ (the upper edge of the mirror) it is necessary to assume $\alpha = 2\pi - \alpha_0 - \theta$ for $F(\theta)$, the lower edge of the mirror, $\alpha = 2\pi - \alpha_0 + \theta$ and for $F(\chi)$

$$\alpha = 2\pi - \frac{\beta_{cr} + \frac{\pi}{2} - \theta}{3}.$$

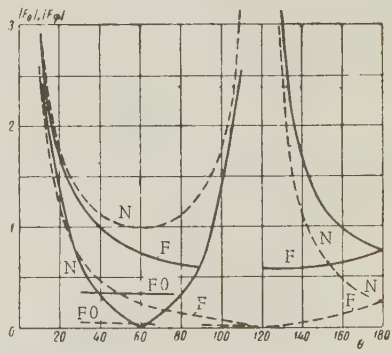


Figure 11

For the θ-component of polarization of lateral radiation F must be replaced with F_0 according to formula (1), with the same values of α, α_0, and χ. The excitation coefficients of the φ and θ components of the fringe wave are

$$\tau_\varphi = (\vec{\varphi}_0 \vec{E}_0),$$

$$\tau_\theta = |[\vec{\varphi}_0 [\vec{E}_0 \vec{\varphi}_0]]|,$$

where $\vec{\varphi}_0$ is the unit vector in the φ direction and E_0 is the unit vector of polarization of the primaty wave at the edge of the mirror. The values of $\tilde{\varepsilon}_1, \tilde{\varepsilon}_2$ are chosen from Tables 1 and 2.

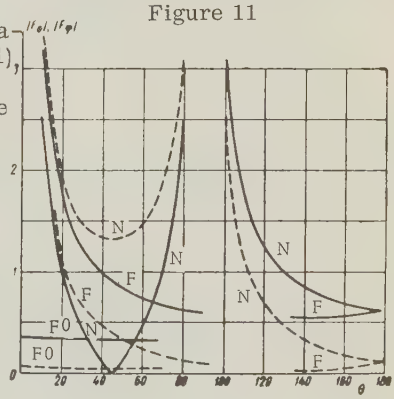


Figure 12

Figures 10-12 present the curves of $F_\varphi, \theta(-\theta)$; $F_\varphi, \theta(\theta)$ and $F_\varphi, \theta(\chi)$ for mirrors of different depth ($\psi_{кр} = 30$; 60; 90°). It follows from the curves that in terms of the φ-component of the field (for example of the pattern in the H-plane) the principal ray for the unshadowed sector proceeds from the near edge of the mirror and the intensity of the ray from the far edge of the mirror is small. The re-reflected ray ($\chi = \lambda\pi$) has small amplitude, since $F_\varphi(2\pi) = 0$.

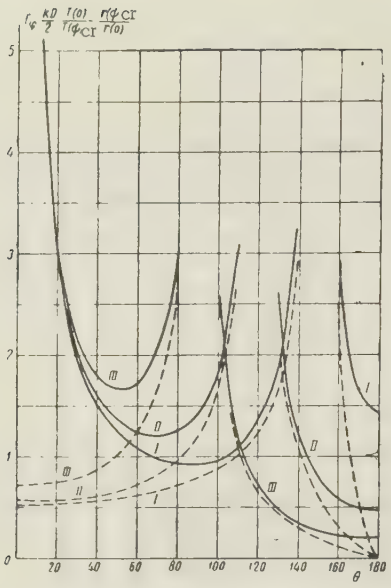


Figure 13. φ-component envelope of lateral radiation for cylindrical mirror (scale along ordinate axis proportional to $\left[\frac{r(\psi_{cr})}{r(0)}\right]^{1/2}$). I, $\psi_{cr} = 30°$; II $\psi_{cr} = 60°$ III, $\psi_{cr} = 90°$; solid line, maximum; dashed line, minimum.

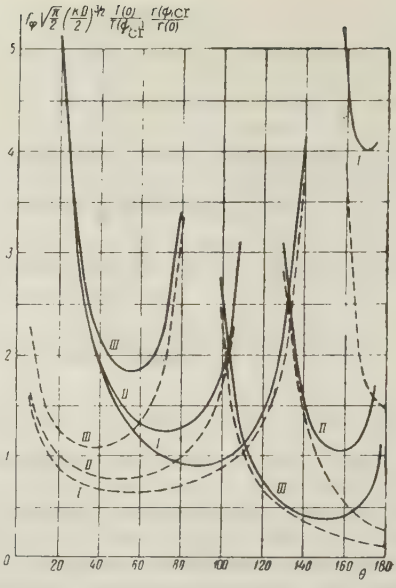


Figure 14. φ-component envelope of lateral radiation for axially symmetric mirror (scale along ordinate axis proportional to $\left[\frac{r(\psi_{cr})}{r(0)}\right]^{1/2}$) I, $\psi_{cr} = 30°$; II, $\psi_{cr} = 60°$ III, $\psi_{cr} = 90°$; solid line, maximum; dashed line, minimum.

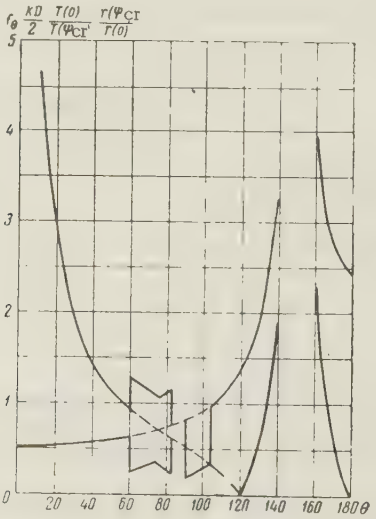


Figure 15. θ-component envelope of pattern of cylindrical mirror; $\psi_{cr} = 30°$ (scale along ordinate axis proportional to $\left[\frac{r(\psi_{cr})}{r(0)}\right]^{1/2}$)

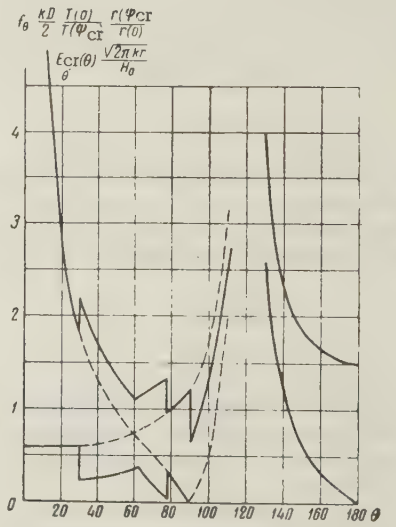


Figure 16. θ-component envelope of pattern of cylindrical mirror; $\psi_{cr} = 60°$

Conversely, for the θ-component of lateral radiation (for example, the pattern in the E-plane) the near edge is of little significance, for $F_\theta(\pi) = 0$.

The chief contribution is made by the far edge and the value of re-reflected ray is relatively large (particularly in the vicinity of the sector $\theta = \frac{\pi}{2} - \frac{\psi_{cr}}{2}$) and determines the interference character of the radiation. The situation is reversed near the edge of the mirror ($\theta = \pi - \psi_{cr}$) and in the shadow sector, wherein the near edge illuminates more intensely (or is generally the only source). Lateral radiation for the φ-component of the field has a more continuous character than for the θ-component. From (21) it follows that the polarization of lateral radiation for each ray is determined by the relationship of the components in the sum

$$\vec{\varphi}_0 f_\varphi (\vec{\varphi}_0 \vec{E}_0) + \vec{\theta}_0 f_\theta | [\vec{\varphi}_0 [\vec{E}_0 \vec{\varphi}_0]] |, \qquad (25)$$

each of which is sinusoidally dependent on φ with period π. The dependence on depth of the mirror ψ_{cr} is monotonic. With an increase in ψ_{cr} the lateral radiation is concentrated in the forward hemisphere.

Figures 13 and 14 present the envelopes of lateral radiation for the φ-component for cylindrical and axially symmetric mirrors. Figures 15-17 show the envelopes of patterns of cylindrical mirrors for the θ-component. The shadow sectors of the radiator and of the fringe waves were taken into account in the calculations. Envelope discontinuities due to shadowing of the fringe waves are mitigated by diffraction processes and depend on kD. Expression (21) is normalized for the maximum of the main lobe. The maximum of the fringe lobe of the radiator is limited by the value of (G_{reg}/G_{ant}), since for it the mirror serves as an ordinary shadow screen. The envelopes of the patterns in the vicinity of the main lobe are proportional to $(\sin\theta)^{-1}$ for cylindrical and $(\sin\theta)^{-3/2}$ for axially symmetric mirrors. The envelopes of the fringe lobes of the radiator in both cases are proportional to $[\sin(\pi-\theta)]^{-1/2}$. Finally, the envelope of the rear lobe ($\theta \sim \pi$) for axially symmetric mirrors is proportional to $\sin(\pi-\theta)^{-1/2}$. The dependence on mirror dimensions and wavelength is $-(kD)^{-1}$ for cylindrical and $(kD)^{-3/2}$ for axially symmetric mirrors.

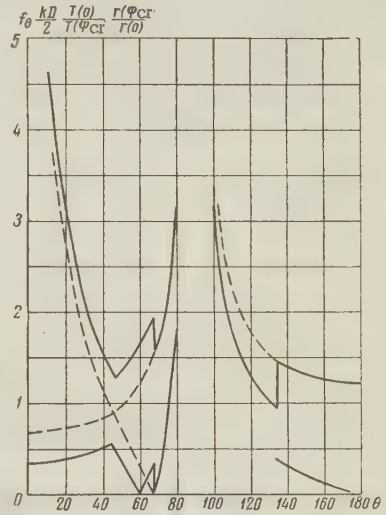


Figure 17. Envelope of pattern of cylindrical mirror for θ-component, $\psi_{cr} = 90°$ (scale along ordinate axis proportional to $\left[\frac{r(\psi_{cr})}{r(0)}\right]^{1/2}$)

The above analysis of envelopes of lateral radiation has dealt with the case of a decrease in illumination at the edge of the mirrors by not more than 10-20 db.

Calculations of the main lobe of the patterns is more easily performed by the aperture method (parabolic mirror) or by means of geometric optics and fringe waves. The region of effective application of the current method is relatively small and limited to sectors of the fringe lobes.

If we disregard the conditions of shadowing of the near field of the radiator, the fringe component of current and multiple reflections, then calculation by the method presented above yields the same results as calculation of the pattern by the current method (see, for example references [7] and [8]). The actual difference in calculations by the two methods taking into consideration the condition of shadowing for the radiator is relatively small and does not exceed 6-8 db, although calculation by the "ray" method is undoubtedly simpler, more descriptive and considers all terms of similar magnitude.

Improvement of the current method must take into account, in particular, the current from multiply reflected rays and the whispering-gallery waves excited by the "fringe" wave. These components are far from small and comprise, for example, for a mirror with a diameter of 100λ from 0.16 at the "edge" to 1.2 at the center for the first component and from 0.22 at the "edge" to 1.5 at the center for the second component relative to the $2[n\vec{H}_0]$ component. In conclusion the author takes pleasure in expressing his thanks to L.A. Vaynshteyn for discussion and advice in preparing this report.

REFERENCES

1. J.B. Keller, Diffraction by an aperture, J. Appl. Phys., 1957, 28, 2, 426.
2. J.B. Keller, Diffraction by convex cylinder, IRE Trans., 1956, AP-4, 3, 312.
3. B.Ye. Kinber, Shadowing condition and diffraction correction for current distribution, Radiotakhnika i elektronika, 1960, 5, 9, 1407.
4. L.A. Vaynshteyn, Difraktsiya elektromagnitnykh i zvukovykh voln na otkrytom kontse volnovoda "Diffraction of electromagnetic and sound waves at the open end of a waveguide", Izd. Sovetskoye radio, 1953.
5. L.A. Vaynshteyn, Elektromagnitnoye pole, "The electromagnetic field", Izd. Sovetskoye radio, 1957.
6. P.Ya. Ufimtsev, Secondary diffraction of electromagnetic waves along a ribbon, ZhTF. 1958, 28, 3, 569.
7. Yu.K. Murav'yev, V.A. Varyukhin, Calculation of lateral and back radiation of antenna in the form of a paraboloid of rotation, Tr. Akademii, 1952, symp. 31, page 117.
8. L.B. Tartakovskiy, Lateral radiation of ideal paraboloid, Radiotekhnika i elektronika, 1959, 4, 6, 920.

Submitted to the editors
11 May 1960

SYMMETRICAL ILLUMINATION OF FINITE BODIES OF ROTATION

P. Ya. Ufintsev

The report is devoted to refinement of the Kirchhoff approximation for finite bodies of rotation with interrupted surface. Refinement is based on approximate calculation of disturbance of the field by the break and is performed for the case of illumination of bodies along their axis of symmetry. The effective scattering surface of a finite cone and paraboloid of rotation is calculated and compared with experimental values.

INTRODUCTION

The problem of diffraction of electromagnetic waves at ideally conducting bodies of finite dimensions with broken surface is of considerable interest but, in virtue of its

complexity, it has not yet been satisfactorily solved. In the region of wavelengths which are short in comparison with the linear dimensions of the test bodies, the Kirchhoff approximation is usually used, but the latter often leads to incorrect results and requires refinement.

In the special case with illumination of convex bodies of rotation along their axis of symmetry we determined refined expressions for the effective scattering surface. The method of calculation has been previously discussed by us in references [1, 2] and consists in determining the scatter field in the form of "uniform" and "nonuniform" components. The uniform component is the scatter field by the Kirchhoff approximation and is determined by integration of the surface current

$$\vec{j} = \frac{c}{2\pi}[\vec{n}\vec{H}],$$

where c is the velocity of light in vacuum, $\vec{n}$ is the external normal to the surface of the body and $\vec{H}$ is the magnetic field of the incident wave. The nonuniform component is the additional field caused by the break of the surface, which must be taken into account in order to obtain the most accurate results. If the linear dimensions of the body are great in comparison with the wavelength, then the surface currents exciting this field and localized in the vicinity of the break may be approximately calculated at each element of the break as if it were the edge of the corresponding dihedral angle.

It is evident that the currents caused by a round break in any convex surface of rotation will be the same as on the corresponding conical body (Figure 1). Hence, the nonuniform component of the field is conveniently studied in the example of this body.

1. NONUNIFORM COMPONENT OF FIELD

Let a plane electromagnetic wave be incident at a conical body in the positive direction of the z-axis (Figure 1). From the equations

$$\left.\begin{aligned} \vec{E} &= -\frac{1}{ik}(\operatorname{grad}\operatorname{div}\vec{A} + k^2\vec{A}, \\ \vec{H} &= \operatorname{rot}\vec{A}, \quad k = \frac{2\pi}{\lambda} \end{aligned}\right\} \tag{1}$$

we find the following expressions for the scatter field in the wave zone

$$\left.\begin{aligned} E_x &= H_y = ikA_x, \\ E_y &= -H_x = ikA_y \end{aligned}\right\} \text{ with } \vartheta = 0 \tag{2}$$

and

$$\left.\begin{aligned} E_x &= -H_y = ikA_x, \\ E_y &= H_x = ikA_y \end{aligned}\right\} \text{ with } \vartheta = \pi. \tag{3}$$

In addition, to the nonuniform component of current flowing in the vicinity of the round break of the conical surface there corresponds the vector potential

$$\vec{A} = \frac{1}{c}\frac{e^{ikr}}{r}\int_0^{2\pi}\left[\int_0^{l_1}\vec{j}_1(\zeta)\,e^{\pm ik\zeta\cos\omega}(a-\zeta\sin\omega)\,d\zeta + \right.$$
$$\left. + \int_0^{l_2}\vec{j}_2(\zeta)\,e^{\mp ik\zeta\cos\Omega}(a-\zeta\sin\Omega)\,d\zeta\right]d\psi$$

or

$$\vec{A} = \frac{a}{c}\frac{e^{ikr}}{r}\int_0^{2\pi}\left[\int_0^{\infty}\vec{j}_1(\zeta)\,e^{\pm ik\zeta\cos\omega}\,d\zeta + \int_0^{\infty}\vec{j}_2(\zeta)\,e^{\mp ik\zeta\cos\Omega}\,d\zeta\right]d\psi, \tag{4}$$

since $\vec{j}_{1(2)}(\zeta) \neq 0$ only in the vicinity of the break ($\zeta = 0$). Here the upper sign in the exponents refers to the case $\vartheta = \pi$; $\vec{j}_1$ is the current flowing to the left of the break; $\vec{j}_2$ is the current flowing to the right of the break; r is the distance from the break to the observation point.

The nonuniform component of current at a sufficiently small element of the conical

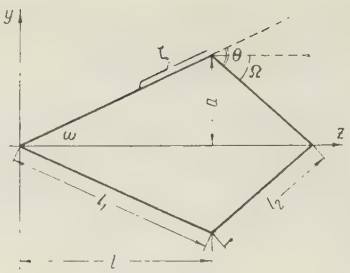


Figure 1.

surface may, in the vicinity of its break, be considered approximately the same as at the corresponding wedge (Figure 2). In the local cylindrical system of coordinates (r_1, φ_1, z_1) the expression for the nonuniform component of the field of such a wedge takes the following form

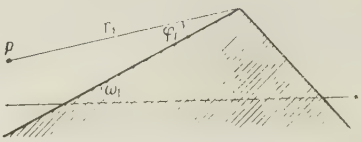


Figure 2.

with $kr_1 \gg 1$:

$$\left.\begin{aligned} E_{z_1}(\psi) &= -H_{\varphi_1}(\psi) = ikA_{z_1}(\psi), \\ H_{z_1}(\psi) &= E_{\varphi_1}(\psi) = ikA_{\varphi_1}(\psi), \end{aligned}\right\} \tag{5}$$

where

$$\vec{A}(\psi) = \frac{1}{c}\sqrt{\frac{2\pi}{kr_1}}\left[\int_0^\infty \vec{j}_1(\zeta)\, e^{\mp ik\zeta\cos\omega}\, d\zeta + \int_0^\infty \vec{j}_2(\zeta)\, e^{\mp ik\zeta\cos\Omega}\, d\zeta\right] e^{i\left(kr_1+\frac{\pi}{4}\right)} \tag{6}$$

Here the upper sign in the exponents refers to the case where $\varphi_1 = \pi + \omega$, and the lower sign refers to the case where $\varphi_1 = \omega$. In reference [1] it was shown that

$$\left.\begin{aligned} E_{z_1}(\psi) &= E_{0z_1}(\psi)\, f \frac{e^{i\left(kr_1+\frac{\pi}{4}\right)}}{\sqrt{2\pi kr_1}}, \\ H_{z_1}(\psi) &= H_{0z_1}(\psi)\, g \frac{e^{i\left(kr_1+\frac{\pi}{4}\right)}}{\sqrt{2\pi kr_1}}, \end{aligned}\right\} \tag{7}$$

where $E_{0z_1}(\psi)$, $H_{0z_1}(\psi)$ are the values of amplitude of the incident wave at the ridge of the wedge; f and g are the angular functions of the nonuniform component of the field scattered by the wedge. Let us introduce the designation

$$\vec{J} = \int_0^\infty (\vec{j}_1(\zeta)\, e^{\pm ik\zeta\cos\omega} + \vec{j}_2(\zeta)\, e^{\mp ik\zeta\cos\Omega})\, d\zeta, \tag{8}$$

then from formulas (5) — (7) it follows that

$$J_{z_1} = \frac{cE_{0z_1}(\psi)}{ik2\pi} f, \quad J_{\varphi_1} = \frac{cH_{0z_1}(\psi)}{ik2\pi} g. \tag{9}$$

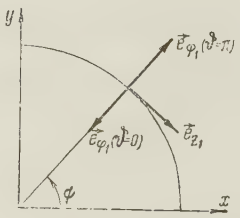


Figure 3

Components J_{z1} and $J_{\varphi 1}$ are mutually perpendicular and parallel to plane xoy (Figure 3). The different orientation of unit vector $\vec{e}_{\varphi 1}$ with $\vartheta = 0$ and $\vartheta = \pi$ is associated with the fact that angle φ_1 is reckoned from the illuminated face of the wedge. In the principal system of coordinates (x, y, z) vector $\vec{J}$ has components

$$\left.\begin{aligned} J_x &= J_{z_1}\sin\psi - J_{\varphi_1}\cos\psi, \\ J_y &= -J_{z_1}\cos\psi - J_{\varphi_1}\sin\psi \end{aligned}\right\} \text{ with } \vartheta = 0 \tag{10}$$

and

$$\left.\begin{aligned} J_x &= J_{z_1}\sin\psi + J_{\varphi_1}\cos\psi, \\ J_y &= -J_{z_1}\cos\psi + J_{\varphi_1}\sin\psi \end{aligned}\right\} \text{ with } \vartheta = \pi. \tag{11}$$

In accordance with formula (9) these expressions may be presented in the form

$$\left.\begin{aligned} J_x &= \frac{c}{ik2\pi}\,[fE_{0z_1}(\psi)\sin\psi - gH_{0z_1}(\psi)\cos\psi], \\ J_y &= -\frac{c}{ik2\pi}\,[fE_{0z_1}(\psi)\cos\psi + gH_{0z_1}(\psi)\sin\psi] \end{aligned}\right\} \text{ with } \vartheta = 0 \tag{12}$$

and

$$\left.\begin{aligned} J_x &= \frac{c}{ik2\pi}\,[fE_{0z_1}(\psi)\sin\psi + gH_{0z_1}(\psi)\cos\psi], \\ J_y &= -\frac{c}{ik2\pi}\,[fE_{0z_1}(\psi)\cos\psi - gH_{0z_1}(\psi)\sin\psi] \end{aligned}\right\} \text{ with } \vartheta = \pi. \tag{13}$$

Now, identifying points in the vicinity of the break in the conical surface with points on the wedge, we obtain in accordance with (4)

$$\left.\begin{aligned} A_x &= \frac{a}{ik2\pi}\frac{e^{ikr}}{r}\int_0^{2\pi}[fE_{0z_1}(\psi)\sin\psi - gH_{0z_1}(\psi)\cos\psi]\,d\psi, \\ A_y &= -\frac{a}{ik2\pi}\frac{e^{ikr}}{r}\int_0^{2\pi}[fE_{0z_1}(\psi)\cos\psi + gH_{0z_1}(\psi)\sin\psi]\,d\psi \end{aligned}\right\} \text{ with } \vartheta = 0, \tag{14}$$

$$\left.\begin{aligned} A_x &= \frac{a}{ik2\pi}\frac{e^{ikr}}{r}\int_0^{2\pi}[fE_{0z_1}(\psi)\sin\psi + gH_{0z_1}(\psi)\cos\psi]\,d\psi, \\ A_y &= -\frac{a}{ik2\pi}\frac{e^{ikr}}{r}\int_0^{2\pi}[fE_{0z_1}(\psi)\cos\psi - gH_{0z_1}(\psi)\sin\psi]\,d\psi \end{aligned}\right\} \text{ with } \vartheta = \pi. \tag{15}$$

Let the incident plane wave be so polarized that $\vec{E} \parallel ox$. Then

$$E_{0z_1}(\psi) = E_{0x}\sin\psi, \quad H_{0z_1}(\psi) = -E_{0x}\cos\psi \tag{16}$$

and, consequently, the nonuniform component of the field caused by the round break will be expressed, in accordance with relationships (2), (3), and (14), (15) as follows:

$$\left.\begin{aligned} E_x &= H_y = \frac{aE_{0x}}{2}(f+g)\frac{e^{ikr}}{r}, \\ E_y &= H_x = 0 \end{aligned}\right\} \text{ with } \vartheta = 0 \tag{17}$$

and

$$\left.\begin{aligned} E_x &= -H_y = -\frac{aE_{0x}}{2}(f-g)\frac{e^{ikr}}{r}, \\ E_y &= H_x = 0 \end{aligned}\right\} \text{ with } \vartheta = \pi. \tag{18}$$

Formula (17) is applicable for values of $0 \leqslant \omega \leqslant \pi/2$, $\omega < \Theta \leqslant \pi$, and formula (18) is applicable for $0 \leqslant \omega \leqslant \pi/2$, $0 \leqslant \Theta \leqslant \pi$. In the case of a disc ($\omega = \pi/2$, $\theta = \pi$), when $f = -g = -1/2$ for $\vartheta = 0$ and $f = g = -1/2$ for $\vartheta = \pi$, the nonuniform component of the field at the z-axis is equal to zero (compare reference [2]).

Using the derived results, let us turn to calculation of the effective scatter surface of specific bodies on the assumption that they are illuminated by a plane wave

$$E_x = E_{0x}e^{i(kz-\omega t)}, \tag{19}$$

and that their linear dimensions are great in comparison with the wavelength.

2. CONE

Let a cone (Figure 1) be illuminated by a plane electromagnetic wave (19). The uniform component of current excited at its surface has the components

$$\left.\begin{aligned} j_x &= \frac{c}{2\pi}E_{0x}\sin\omega\, e^{ikz}, \\ j_y &= 0, \\ j_z &= \frac{c}{2\pi}E_{0x}\cos\omega\cos\psi e^{ikz} \end{aligned}\right\} \tag{20}$$

and creates in the direction $\vartheta = \pi$ a field

$$\left.\begin{aligned} E_x = -H_y = -E_{0x}\frac{i}{4k}\operatorname{tg}^2\omega\frac{e^{ikR}}{R} + E_{0x}\left(\frac{i}{4k}\operatorname{tg}^2\omega + \frac{a}{2}\operatorname{tg}\omega\right)\frac{e^{ikR}}{R}e^{2ikl}, \\ E_y = H_x = 0. \end{aligned}\right\} \tag{21}$$

The nonuniform component of the field, caused by a break of the surface at the base of the cone, is defined, in accordance with (18), by the expression

$$\left.\begin{aligned} E_x = -H_y = -\frac{aE_{0x}}{2}\left(\operatorname{tg}\omega + \frac{\frac{2}{n}\sin\frac{\pi}{n}}{\cos\frac{\pi}{n} - \cos\frac{2\omega}{n}}\right)\frac{e^{ikR}}{R}e^{2ikl}, \\ E_y = H_x = 0, \end{aligned}\right\} \tag{22}$$

where

$$n = 1 + \frac{\omega + \Omega}{\pi}. \tag{23}$$

Asymptotic evaluation of the exact diffraction series for a semi-infinite cone (reference [3]) shows that in the direction $\vartheta = \pi$ we may neglect the nonuniform component of the field caused by the conical tip. Hence, adding (21) and (22), we obtain the approximate expression for the scatter field:

$$E_x = -H_y = -E_{0x}\left(\frac{1}{2k}\operatorname{tg}^2\omega\sin kle^{ikl} + \frac{a}{2}\frac{\frac{2}{n}\sin\frac{\pi}{n}}{\cos\frac{\pi}{n} - \cos\frac{2\omega}{n}}e^{2ikl}\right)\frac{e^{ikR}}{R}. \tag{24}$$

The effective scatter surface is defined by the formula

$$\sigma = \pi a^2 |\Sigma|^2, \tag{25}$$

where the function Σ is associated with the scatter field by the relationship

$$E_x = -H_y = -\frac{aE_{0x}}{2}\frac{e^{ikR}}{R}\Sigma \tag{26}$$

and is equal to

$$\Sigma = \frac{1}{ka}\operatorname{tg}^2\omega\sin kle^{ikl} + \frac{\frac{2}{n}\sin\frac{\pi}{n}}{\cos\frac{\pi}{n} - \cos\frac{2\omega}{n}}e^{2ikl}. \tag{27}$$

The analogous function in the Kirchhoff approximation may be written in accordance with (21) in the form

$$\Sigma_k = \frac{1}{ka}\operatorname{tg}^2\omega\sin kle^{ikl} - \operatorname{tg}\omega e^{2ikl}. \tag{28}$$

With deformation of the forward portion of the cone into a disc ($\omega \to \pi/2$, $l \to 0$, $\Omega = \text{const}$) formulas (27) and (28) are transformed respectively to the form

$$\left.\begin{aligned} \Sigma = -ika - \frac{1}{n}\operatorname{ctg}\frac{\pi}{n}, \\ \Sigma_k = -ika. \end{aligned}\right\} \tag{29}$$

From (27) and (28) it further follows that for large values of parameter ka($ka \gg \operatorname{tg}^2\omega$)

$$\Sigma = \frac{\frac{2}{n}\sin\frac{\pi}{n}}{\cos\frac{\pi}{n} - \cos\frac{2\omega}{n}}e^{2ikl}, \tag{30}$$

$$\Sigma_k = \operatorname{tg}\omega e^{2ikl}. \tag{31}$$

Thus, even in the case of short waves ($ka \gg \operatorname{tg}^2\omega$, but $R \gg kl^2$) our expression (27) does not become the formula in physical optics and is considerably different from it:

$$\sigma = \pi a^2\left|\frac{\frac{2}{n}\sin\frac{\pi}{n}}{\cos\frac{\pi}{n} - \cos\frac{2\omega}{n}}\right|^2, \tag{32}$$

$$\sigma_k = \pi a^2\operatorname{tg}^2\omega, \tag{33}$$

whence

$$\sigma = \sigma_k \left| \frac{\frac{2}{n} \sin \frac{\pi}{n}}{\left(\cos \frac{\pi}{n} - \cos \frac{2\omega}{n} \right) \operatorname{tg} \omega} \right|^2, \tag{34}$$

that is, for sufficiently short waves (or for sufficiently large cone dimensions) the function of σ is proportional to σ_k, wherein the proportionality factor does not depend on the cone dimensions but depends only on the shape of the cone.

This result is graphically illustrated by calculations of the effective scatter surface of the cone ($\omega = 10°25'$, $k = \pi$, $\Omega = 90°$) as a function of its length (Figure 4). Whereas our formula (unbroken line) is in satisfactory agreement with experiment (isolated points x), the Kirchhoff approximation (dashed line) yields values 13–15 db less than the experimental values. The nonuniform component of the field for sharp cones has a particularly large value. For example, see Figure 5, in which the curve for the effective surface of the cone is plotted ($a = 2.75$ cm, $k = \pi$, $\Omega = 90°$) with its deformation into a disc ($\omega \to 90°$). Departure between our curve and the Kirchhoff approximation is here (at $\omega \simeq 2°$) almost 30 db.

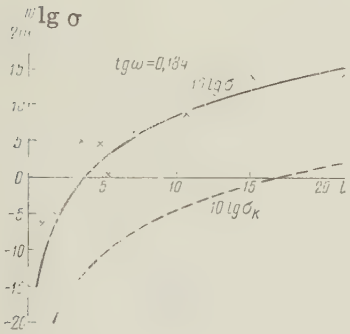


Figure 4

Let it be noted that our expression (32) is equivalent to that derived in reference [4]; however, the latter is applicable only for sharp cones, whereas, we have, in addition to (32), formula (27) which is suitable for cones with any angle of flare ($0 \leqslant \omega \leqslant \pi/2$).

As distinct from the Kirchhoff approximation (28), our expression (27) also permits evaluation of the role of shape of the portion of the body and shows that the closer this shape is to that of a funnel ($\Omega \simeq \pi - \omega$), the larger the reflected signal. For example, in the case $\omega = 10°$, $kl = 10\pi$ ($k = \pi$) the value of the signal reflected by the cone may exceed by 15 db the corresponding Kirchhoff approximation (see Figure 6) if $\Omega \simeq 170°$.

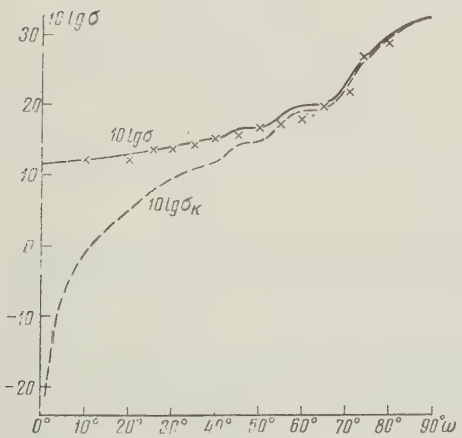


Figure 5

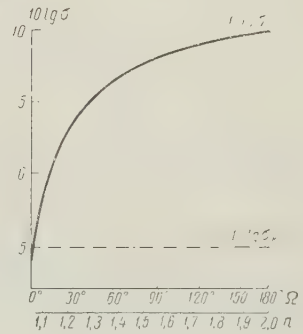


Figure 6

3. PARABOLOID OF REVOLUTION ($r^2 = 2pz$)

The uniform component of current excited at the surface of a paraboloid (Figure 7) by the incident wave (19),

$$\left.\begin{aligned} j_x &= \frac{c}{2\pi} E_{0x} \sin\alpha e^{ikz}, \\ j_y &= 0, \\ j_z &= \frac{c}{2\pi} E_{0x} \cos\alpha \cos\psi e^{ikz} \end{aligned}\right\} \tag{35}$$

creates in the direction $\vartheta = \pi$ a field

$$\left.\begin{aligned} E_x &= -H_y = -\frac{aE_{0x}}{2}(1 - e^{2ikz}) \operatorname{tg}\omega \frac{e^{ikR}}{R}, \\ E_y &= H_x = 0, \end{aligned}\right\} \tag{36}$$

where a is the base radius of the paraboloid, $l = a^2/2p = a/2 \operatorname{ctg}\omega$ is its length, α is the angle formed by the z-axis and the tangent to the parabola $r^2 = 2pz$; at point $z = l$ it assumes the value $\alpha = \omega$ ($\operatorname{tg}\omega = p/a$).

The nonuniform component caused by the circular break in the paraboloid surface is defined by formula (22). The nonuniform component caused by smooth bending of the paraboloid surface is equal to zero in the presence of symmetrical illumination [5]. Hence, adding (36) and (32), we find the expression for the resulting scatter field

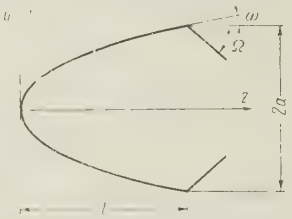


Figure 7

$$E_x = -H_y = -\frac{aE_{0x}}{2}\left(\operatorname{tg}\omega + \frac{\frac{2}{n}\sin\frac{\pi}{n}}{\cos\frac{\pi}{n} - \cos\frac{2\omega}{n}} e^{2ikl}\right)\frac{e^{ikR}}{R}. \tag{37}$$

Consequently, the effective scatter surface of the paraboloid will be defined by the relationship

$$\sigma = \pi a^2 \left| \operatorname{tg}\omega + \frac{\frac{2}{n}\sin\frac{\pi}{n}}{\cos\frac{\pi}{n} - \cos\frac{2\omega}{n}} e^{2ikl} \right|^2, \tag{38}$$

which upon deformation of the paraboloid into a disc ($\omega \to \pi/2$, $l \to 0$, $\Omega = \text{const}$) assumes the form

$$\sigma = \pi a^2 \left| ika + \frac{1}{n} \operatorname{ctg}\frac{\pi}{n} \right|^2. \tag{39}$$

Comparing expression (38) with the formula

$$\sigma_k = \pi a^2 \operatorname{tg}^2\omega \left| 1 - e^{2ikl} \right|^2, \tag{40}$$

which gives the Kirchhoff approximation for the effective scatter surface, we see that they differ considerably. In fact, we note first of all the oscillatory character of the function σ_k: the reflected signal is equal to zero if the length of the paraboloid is equal to a whole number of half waves ($l = \frac{\lambda}{2}$ n, n = 1, 2, 3, ...), and assumes maximum value when it includes half number of half waves ($l = \frac{\lambda}{2}(n + \frac{1}{2})$, n = 1, 2, 3...).

Calculations performed by us with formula (38) for paraboloids with parameters $\Omega = 90°$, $\operatorname{tg}\omega = 0.1$ ($k = \pi$) show (Figure 8) that, while the oscillatory character of the effective

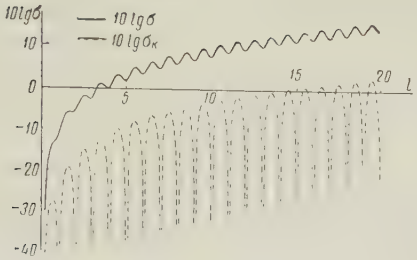


Figure 8

scatter surface is preserved, the amplitude of oscillation is approximately only 2 db and the maximum values of function σ exceed the corresponding values in the Kirchhoff approximation by almost 13 db. Departure between the results of our theory and the Kirchhoff approximation is even more evident in the case of deformation of the paraboloid into a disc (Figure 9, $ka = 3\pi$, $k = \pi$, $\Omega = 90°$, $\omega \to 90°$).

As in the case of the cone, the shape of the shadow portion has considerable effect on the value of the reflected signal. For example, for a paraboloid with parameters $ka = 2\pi$, $kl = 10\pi$, $tg\omega = 0.1$ ($k = \pi$) the value of the reflected signal increases with an increase in angle Ω ($\omega < \Omega < \pi - \omega$) by 44 db (Figure 10)!

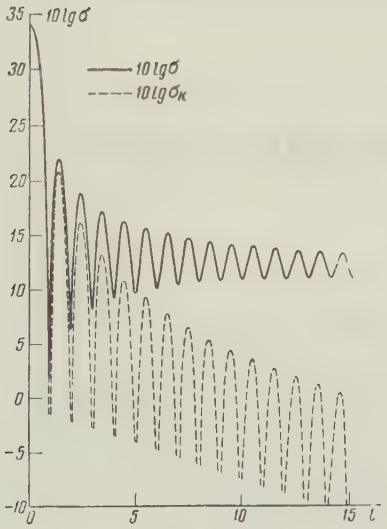


Figure 9

CONCLUSION

The effective scatter surface of a finite cone and of a paraboloid of rotation have been calculated. The linear dimensions of the bodies are assumed to be small in comparison with the wavelength and their surfaces are assumed to be ideally conducting. Illumination by electromagnetic waves is achieved along the axis of symmetry. The absolute system of Gaussian units is used.

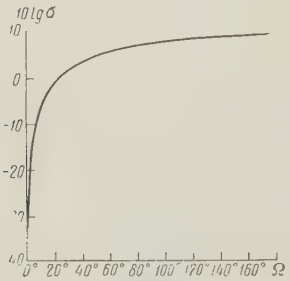


Figure 10

The derived expressions are suitable both for convex and for concave bodies, but in the latter case it is required that the concave portions of the surface lie within the shadow region.

The calculations reveal that the value of the reflected signal is considerably dependent on the shape of the shadow portion of the body and increases with the concavity of shadow portion. However, since the nonuniform component is caused by surface elements adjoining the break, that portion of the shadowed surface which is removed from the break by several wavelengths will have no effect on the value of reflected signal and may be arbitrary.

It is noteworthy that our expressions, which are in satisfactory agreement with experiment, even for large (in comparison with λ) dimensions of the bodies, do not turn into formulas of physical optics but differ considerably from them. At the same time, the Kirchhoff approximation (notwithstanding the current opinion of its reliability in such cases) leads to considerable departure from experimental values.

The results presented in this report permit calculation of the effective scatter surface in the presence of symmetrical illumination of any convex body of rotation the surface of which has circular breaks.

I wish to take this opportunity to express my thanks to Ye.N. Mayzel's and L.S. Chugunova for providing experimental data.

REFERENCES

1. P.Ya. Ufimtsev, ZhTF, 1957, 27, 8, 1840.
2. P.Ya. Ufimtsev, ZhTF, 1958, 28, 11, 2604.
3. L.B. Felsen, J. Appl. Phys., 1955, 26, 2, 138.
4. K.M. Siegel, R.F. Goodrich, V.H. Weston, Appl. Scient. Res. B, 1959, 8, page 8.
5. C.E. Schensted, J.Appl. Phys., 1955, 26, 3, 306.

Submitted to the editors
28 April 1960

RESONANCE TRANSFORMS AND THEIR PROPERTIES

S. Dyad'kov

Resonance transforms of the type $D\omega(t) = \int f'(\tau)e^{-j\omega\tau}d\tau$ and $F_{\omega}(t) = \int f(\tau)e^{-j\omega\tau}d\tau$ as well as their images in the complex plane are investigated in this report. These transforms are closely related to the spectral functions and, particularly, to the spectral density of the current spectrum. In addition, as is shown in the report, these transforms may be conveniently used in the representation of the oscillatory process in a lossless tank circuit to the input of which there is applied a voltage equal to the investigated function $f(\tau)$ and the natural frequency of which is equal to the transformation frequency ω. This is the author's basis for referring to these conversions as "resonance" transforms. Notwithstanding the fact that resonance transforms are closely related to regions which have been thoroughly studied, their properties are as yet unknown to a wide circle of scientific and technical workers and have rarely been dealt with in the literature [3,5]. This report presents a number of theorems dealing with the properties of resonance transforms and their association with the spectra and oscillatory transients. Also presented are considerations dealing with the possibility of extending the resonance transform method to lossy circuits and to circuits with a nonlinear resistance. Finally, several applications of the resonance transform method for the solution of transient processes are listed.

1. DEFINITION OF BASIC CONCEPTS

As undamped resonance-transformed functions of the investigated process expressed as a function of time $u = f(\tau)$ let us designate the expressions

$$D_{\omega}(t) = \int_{0-}^{t} f'(\tau)\, e^{-j\omega\tau}d\tau, \tag{1}$$

$$F_{\omega}(t) = \int_{0}^{t} f(\tau)\, e^{-j\omega\tau}d\tau. \tag{2}$$

The curves representing these functions in the complex plane will be referred to as undampe resonance images and the mathematical operations defined by expressions (1) and (2) will be referred to as undamped resonance transforms. For refinement in certain cases operation (1) will be referred to as a derived, undamped resonance transform of function $f(\tau)$ while operation (2) will be referred to as the principal undamped transforms. For the sake of brevity we will often use abbreviated designations, for example: transformed $D_\omega(t)$, image $D_\omega(t)$, etc. It must be emphasized that transformed $D_\omega(t)$ and $F_\omega(t)$ are functions of time t as the upper limit of integrals (1) and (2), which functions are determinate for a given value of angular frequency ω of transformation. As shown below, $D_\omega(t)$ and $F_\omega(t)$ are, numericall the variable complex amplitudes of voltages at the capacitor and inductor in a series resonant undamped circuit with natural angular frequency ω_K equal to the angular frequency of transformation ω, to the input of which there is applied a voltage $u = f(\tau)$. It was this circumstance which permitted the author to refer to (1) and (2) as undamped resonance transforms.

The notion of resonance transforms may be extended also to damped circuits. In this case we shall be dealing with damped resonance transforms.

In this report we shall limit discussion to those functions $f(\tau)$ which may represent real physical processes, that is, which automatically satisfy the Dirichlet conditions.

Duration of the real process $f(\tau)$ examined by us will be considered finite (with the exception of section 8); the moment of commencement of the process will be represented as $\tau_H \geqslant 0$ and the moment of termination of the process as τ_K. Hence $f(\tau) = 0$ for all $\tau \leqslant 0$ and all $\tau \geqslant \tau_K$.

2. METHOD OF PLOTTING AND GEOMETRIC PROPERTIES OF RESONANCE IMAGES

In order to obtain the image of (1) in the complex plane it is necessary to lay off in succession infinitely small increments of $f'(\tau)d\tau$ [and in the case of jumps in $f(\tau)$, finite] in directions given by the corresponding factors of $e^{-j\omega\tau}$, that is, at an angle

$$\varphi = -\omega\tau \tag{3}$$

to the real coordinate axis. The method of mechanization of this integration has been describe in reference [1].

Since the direction of an infinitely small segment of straight line coincides with the direction of the tangent, then from (3) there follows a tangent theorem which is valid for all resonance images (principal and derived, undamped and damped.)

Theorem 1. The direction of the tangent to the resonance image of function $f(\tau)$ at the point corresponding to moment τ is given by the angle $\varphi = -\omega\tau$ and does not depend on the form of function $f(\tau)$.

Factor $e^{-j\omega\tau}$ in expression (1) defines only the direction of the elementary segments $f'(\tau)d\tau$ in the complex plane but not their absolute value. Hence the length of the resonance image from the origin to the point corresponding to moment t is given by the expression

$$L_d = \int_0^t |f'(\tau)|\, d\tau. \tag{4}$$

From this there follows the theorem on constancy of image length.

Theorem 2. The instantaneous value of image length $D_\omega(t)$ from the origin to moment t is equal to the sum of the absolute values of all increments (finite and infinitely small) of function $f(\tau)$ in the interval from 0 to t and does not depend on the integration frequency ω. The total image length $L_d(\tau_K)$ is equal to the sum of the absolute values of all the differences between values of $f(\tau)$ in adjacent extreme (including the origin and end) and does not vary with a change in ω.

In a similar manner we may determine the length for the principal image of $F_\omega(t)$, only in this case we shall add the absolute values of the elementary increments in area between the curve of $f(\tau)$ and axis τ:

$$L(t) = \int_0^t |f(\tau)|\, d\tau, \tag{5}$$

whence we have

Theorem 3. The instantaneous value of length of the image of $F_\omega(t)$ is equal to the sum of absolute values of the positive and negative sectors of the area between curve $f(\tau)$, axis τ and the ordinates at points $\tau = 0$ and $\tau = t$.

The radius of curvature is equal to the ratio of an infinitely small increment of arc to the corresponding angle between infinitely close tangents. Consequently, for the image $D_\omega(t)$ we may write

$$\rho_d(t) = \frac{|f'(t)|\,d\tau}{d\varphi} = \frac{|f'(t)|\,d\tau}{-\omega d\tau} = -\frac{|f'(t)|}{\omega}, \tag{6}$$

and for the image $F_\omega(t)$

$$\rho(t) = \frac{|f(t)|\,d\tau}{-\omega d\tau} = -\frac{|f(t)|}{\omega}. \tag{7}$$

Whence we obtain

Theorem 4. The radius of curvature of image $D_\omega(t)$ (or $F_\omega(t)$) at the point corresponding to moment t is equal to the value of the derivative of f'(t) (or the function f(t) itself) divided by angular frequency ω and does not depend on the residual values of function $f(\tau)$.

Conclusions. For each linear segment of $f(\tau)$ the derivative of $f'(\tau)$ is constant and hence the corresponding portion of the image $D_\omega(t)$ is a circle of radius

$$r_d = \frac{f'(t)}{\omega}. \tag{8}$$

If, moreover, $f'(\tau) = \infty$, which corresponds to a jump in function $f(\tau)$m then $r_d = \infty$, i.e., the jumps in function $f(\tau)$ are reflected as linear segments $D_\omega(t)$. If $f'(\tau) = 0$, i.e., in the given segment $f(\tau)$ is constant, then $r_d = 0$, i.e., this segment is imaged as a point and the angle between adjacent tangents is proportional to the length of the sector.

At the points of the extrema of function $f(\tau)$ the derivative of $f'(\tau)$ changes its sign and at these points the image $D_\omega(t)$ has a cusp, since here the direction in which it is necessary to lay off the elementary increments is reversed.

For each sector where $f(\tau)$ is constant the radius of curvature of image $F_\omega(t)$ is constant and equals

$$r = \frac{f(t)}{\omega}. \tag{9}$$

At points where $f(\tau)$ changes sign the image $F_\omega(t)$ has a cusp.

Let us determine the relationship between images $D_\omega(t)$ and $F_\omega(t)$. We shall use the formula for integration by parts for $F_\omega(t)$:

$$F_\omega(t) = \int_0^t f(\tau)\,e^{-j\omega\tau}d\tau = \frac{1}{j\omega}\left(\int_0^t f'(\tau)\,e^{-j\omega\tau}d\tau - f(t)\,e^{-j\omega t}\right)$$

or

$$F_\omega(t) = \frac{1}{j\omega}\left(D_\omega(t) - f(t)\,e^{-j\omega t}\right). \tag{10}$$

The last term in parentheses indicates that at point t it is necessary to add to image $D_\omega(t)$ in the negative direction segment f(t) at angle $\varphi = -\omega t$. But at the same angle the corresponding element of arc of image $D_\omega(t)$ is laid off. Consequently, segment f(t) must have a direction tangent to $D_\omega(t)$. Moreover, on the basis of Theorem 2, the increment in length of arc of $D_\omega(t)$ is equal to the increment of segment f(t). Consequently, the curve of $D_\omega(t) - f(t)e^{-j\omega t}$ is obtained by transformation of a straight line for $D_\omega(t)$, that is, it is the involute with respect to $D_\omega(t)$. Hence, we obtain

Theorem 5. The image of $F_\omega(t)$ is the involute of the image of $D_\omega(t)$ divided by $j\omega$.

3. EXAMPLES OF RESONANCE TRANSFORMS

On the basis of the theorems proved in the previous section it is easy to plot resonance images for many forms of $f(\tau)$. For example, for the straight line $f(\tau) = a(\tau)$ the derivative $f'(\tau) = a$ and hence $D_\omega(t)$ will be imaged, in accordance with (8), as a circle of radius

$r_d = a/\omega$ with center at point $(0, -a/\omega)$(Figure 1.) On the basis of (10) $F_\omega(t)$ will be imaged as the involute of a circle of radius $r = a/\omega^2$ with center at $(-a/\omega^2, 0)$. For various values of ω we obtain a family of circles and involutes.

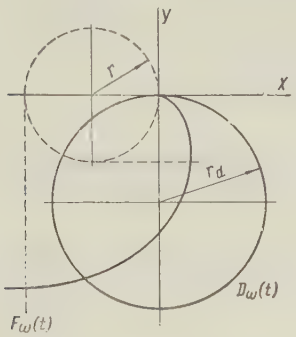


Figure 1. Transformed $D_\omega(t)$ and $F_\omega(t)$ of straight line $f(\tau) = a\tau$.

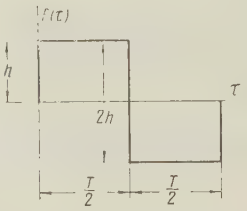


Figure 2. Function f(t) in the form of a single period of square wave.

For the case where $f(\tau)$ is given by a single period of a square wave (Figure 2) the images of $D_\omega(t)$ (Figure 3) are obtained in the form of a broken line consisting of three segments: h, 2h, and h, while the images of $F_\omega(t)$ are obtained in the form of arcs of two circles (Figure 4) of radius $r=h/\omega$ in accordance with (7). Transition of $f(\tau)$ from positive to negative values is represented in the image of $F_\omega(t)$ as a cusp formed by two arcs.

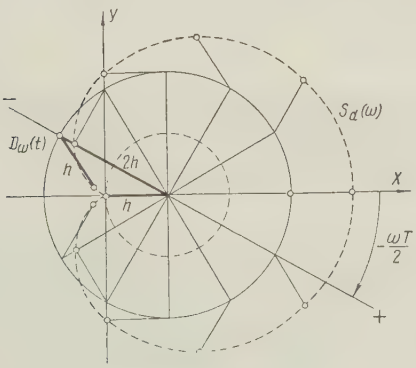


Figure 3. Family of transformed $D_\omega(t)$ of function $f(\tau)$ imaged in Figure 2; $S_d(\omega)$ is the spectrum of the derivative of this function.

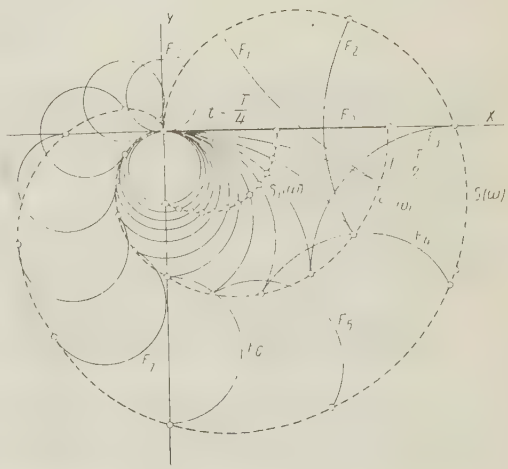


Figure 4. Family of transformed $F_\omega(t)$ of function $f(\tau)$ imaged in Figure 2 for $\omega = 0$. $\Omega/6$, $\Omega/3$, $\Omega/2, \ldots, 2\Omega$, where $\Omega = 2\pi/T$ is the angular frequency of the square waves (Figure 2); $S(\omega)$ is part of the complex spectrum from $\omega = 0$ to $\omega = 2\Omega$; $S_1(\omega)$ and $S_2(\omega)$ are curves of the current spectrum of $t = T/4$ and $t = T/2$.

For a sawtooth wave the images of $D_\omega(t)$ consist of segments of straight lines Figure (5) corresponding to jumps and of arcs of circles corresponding to a linear rise in $F(\tau)$. The transformed $F_\omega(t)$ in this case is imaged by arcs of involutes of circles. In reference [1] it was shown that the image of $D_\omega(t)$ for $f(\tau) = \sin \omega_1 \tau$ is a cycloid (Figure 6) for $\omega = \omega_1$ and a hypocycloid or epicycloid for $\omega \neq \omega_1$.

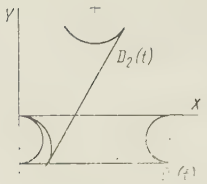


Figure 5. Examples of transformed $D_\omega(t)$ of a single period of sawtooth wave.

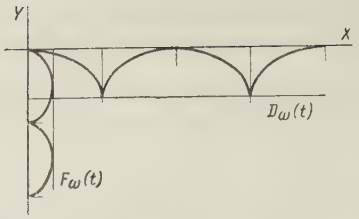


Figure 6. Cycloids imaging transformed $D_\omega(t)$ and $F_\omega(t)$ of a single period of sine wave.

4. RELATION BETWEEN UNDAMPED RESONANCE TRANSFORMS OF FUNCTION f(τ) AND ITS SPECTRUM

Keeping in mind that (as was stated in section 1) $f(\tau) = 0$ for all $\tau \leqslant 0$ and all $\tau \geqslant \tau_K$ we may rewrite expressions (1) and (2) as follows:

$$D_\omega(\tau_K) = \int_0^{\tau_K} f'(\tau)\, e^{-j\omega\tau} d\tau = \int_{-\infty}^{+\infty} f'(\tau)\, e^{-j\omega\tau} d\tau = S_d(\omega), \tag{11}$$

$$F_\omega(\tau_K) = \int_0^{\tau_K} f(\tau)\, e^{-j\omega\tau} d\tau = \int_{-\infty}^{+\infty} f(\tau)\, e^{-j\omega\tau} d\tau = S(\omega). \tag{12}$$

This means that the values of transformed $D_\omega(t)$ or $F_\omega(t)$ calculated for moment $t = \tau_K$ are numerically equal to the spectral density $S_d(\omega)$ of the derivative of function $f'(\tau)$ or the spectral density $S(\omega)$ of function $f(\tau)$ itself for given ω. The values of transforms calculated for moment t yield values of the current spectral density for this moment

$$D_\omega(t) = S_{dt}(\omega), \tag{13}$$

$$F_\omega(t) = S_t(\omega). \tag{14}$$

Hence, it follows that if in the family of curves (see, for example, Figure 3) imaging the transformed $D_\omega(t)$ for different values of ω we join the terminal points of $D_\omega(\tau_K)$, then we obtain a curve reflecting the complex spectrum of function $f'(\tau)$. If, in the family of $D_\omega(t)$ we join the points of $D_\omega(t)$ corresponding to moment t, we obtain the curve of the current spectrum for this moment t.

In similar manner, by joining in the family of curves of $F_\omega(t)$ e.g., Figure 4 the end points of $F_\omega(\tau_K)$, we obtain the curve of the complex spectrum of $s(\omega)$ of function $f(\tau)$ and, by joining the points of $F_\omega(t)$, we obtain the value of its current spectrum $S_t(\omega)$ for moment t.

5. DETERMINING VOLTAGE AT INDUCTOR AND CURRENT IN A LOSSLESS SERIES RESONANT CIRCUIT BY MEANS OF RESONANCE TRANSFORM $D_\omega(t)$

It is customary to express the instantaneous value a(t) of a harmonically varying quantity as the projection of a uniformly rotating, constant-valued vector of amplitude A. This method of representation may be described as follows

$$a(t) = \mathrm{Re}\, Ae^{j\omega t} \tag{15}$$

The representation of any variable (anharmonic) quantity as the projection onto the real axis of an amplitude vector of variable magnitude and phase is not single-valued, For single-valued representation it is necessary to introduce an additional condition. Such a condition may be the requirement that this vector be represented as the projection onto an imaginary axis of a second given variable b(t). Then the two time functions representing the given quantities a(t) and b(t) uniquely define the time changes in absolute

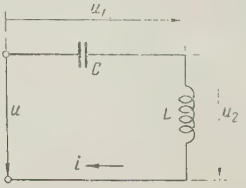


Figure 7. Lossless circuit with input voltage $u = f(\tau)$.

magnitude and phase of the variable amplitude vector. As shown below, definition of the inductor voltage $u_2(t)$ and current i(t) in a tuned circuit (Figure 7) by means of the resonance transform of $D_\omega(t)$ permits expressing the current and voltage by one vector, which presents considerable advantages, particularly in the design of devices operating on the resonance transform principle [1,2].

Let us assume that to the input of a lossless series resonant circuit (Figure 7) there is applied a voltage $u = f(\tau)$ numerically equal to the values of the processes under investigation. In order to determine the instantaneous values $U_2(t)$ of the voltage at the inductor let us use the Duhamel integral

$$u_2(t) = f(0)\,h(t) + \int_0^t f'(\tau)\,h(t-\tau)\,d\tau, \tag{16}$$

where h(t) is the transfer function and f(0) is the value of $f(\tau)$ when $\tau = 0$, which, in accordance with section 1 is equal to zero. For our case the transfer function may be represented by the following expressions

$$h(t) = \frac{e^{j\omega_K t} + e^{-j\omega_K t}}{2}. \tag{17}$$

where $\omega_K = 1/\sqrt{LC}$ is the natural frequency of the tuned circuit.

Inserting this expression into (16), we obtain

$$u_2(t) = \mathrm{Re}\left[\int_0^t f'(\tau)\,e^{-j\omega_K \tau}d\tau\right] e^{j\omega_K t}. \tag{18}$$

Comparing this expression with (1), we see that the quantity in parentheses is the instantaneous value of transformed $D_\omega(t)$ with $\omega = \omega_K$, so that we may write

$$u_2(t) = \mathrm{Re}\, D_\omega(t)\, e^{j\omega t}\,|_{\omega=\omega_K}. \tag{19}$$

Comparing this expression with (15), we conclude that D(t) may be regarded as the complex variable amplitude $U_2(t)$ of the voltage at the inductor:

$$U_2(t) = \int_0^t f'(\tau)\,e^{-j\omega_K \tau}d\tau = D_\omega(t)\,|_{\omega=\omega_K}. \tag{20}$$

The resulting relationship may be formulated as follows.

Theorem 6. The variable complex amplitude of voltage at the inductor in a lossless series circuit with input voltage $u = f(\tau)$ may be represented as a transformed $D_\omega(t)$ if it is assumed that the angular frequency of transformation is equal to the natural angular frequency of the tuned circuit.

Solving the Duhamel integral for current,

$$i(t) = \int_0^t f'(\tau)\,A(t-\tau)\,d\tau,$$

where the transfer admittance A(t) may be represented by the expression

$$A(t) = \frac{1}{2j\omega L}\left(e^{j\omega_K t} - e^{-j\omega_K t}\right),$$

we obtain

$$i(t) = \mathrm{Re}\left[\frac{1}{j\omega L}\int_0^t f'(\tau)\,e^{-j\omega_K \tau}d\tau\right] e^{j\omega_K t},$$

whence, on the basis of (1),

$$i(t) = \mathrm{Re}\,\frac{D_\omega(t)}{j\omega L}\,e^{j\omega t}\,|_{\omega=\omega_K}. \tag{21}$$

Consequently, on the basis of (15), the variable amplitude of current may be represented by the complex vector

$$I(t) = \frac{1}{j\omega L}\int_0^t f'(\tau)\,e^{-j\omega\tau}d\tau = \frac{D_\omega(t)}{j\omega L}\Big|_{\omega=\omega_K}. \tag{22}$$

From this we obtain

Theorem 7. The variable complex amplitude in a lossless series circuit to the input of which there is applied a voltage $u = f(\tau)$ may be represented by the transformed $D_\omega(t)$ divided by $j\omega L$ if it is assumed that the angular frequency of transformation is equal to the natural angular frequency of the circuit.

Conclusion. From (20) and (22) it follows that

$$\frac{U_2(t)}{I(t)} = j\omega L,$$

that is, that the ratio of the variable amplitudes of voltage at the inductor and the current, determined by means of the transformation of $D_\omega(t)$, is the same as for constant amplitudes in the presence of harmonic steady-state processes.

It follows from (21) that

$$i(t)\,\omega L = \mathrm{Im}\,D_\omega(t)\,e^{j\omega t}\,|_{\omega=\omega_K}, \tag{23}$$

i.e., that the values of current (on a scale determined by factor ωL) may be represented as projections of vector $D_\omega(t)e^{j\omega\tau}$ onto an imaginary axis. Comparing this result with (19), we may formulate the following theorem.

Theorem 8. The value of transformed $D_\omega(t)$ wholly defines the state of the oscillatory process for the moment t, for the current and the voltage at the inductor are simultaneous defined.

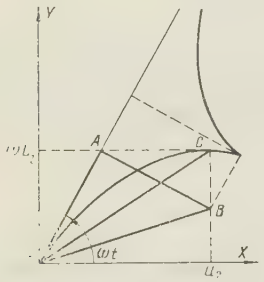


Figure 8. Representation of current i(t) and voltage $u_2(t)$ in lossless circuit by means of transformed $D_\omega(t)$ of input voltage.

In this respect, the method of transformation of $D_\omega(t)$ is similar to the phase density method, wherein the position of a point defines the state of the oscillatory system. However, the derived representations are substantially different and the representation $D\omega(t)$ more graphically illustrates the effect of the applied voltage $u = f(\tau)$

It must be pointed out that the generally accepted methods of solution of transients do not yield clearly expressed values of variable amplitudes satisfyin the possibilities of the image of $u_2(t)$ and i(t) as a single rotating vector. Thus, in solving by the operator method the transient process in an undamped circuit upon the application of a sinusoidal voltage $u(t) = U \sin \omega t$ for $\omega = \omega_K$, we obtain the expression for current

$$i(t) = \frac{U}{2L}\,t \sin \omega t \tag{24}$$

and for voltage at the inductor

$$u_2(t) = \frac{U\omega t}{2}\cos \omega t + \frac{U}{2}\sin \omega t.$$

The quantity $\omega Li(t)$, reflecting on a given scale the current i(t), may be represented as the projection onto the Y-axis of a uniformly rotating vector $OA = U\omega t/2$, rotating with angular frequency ω (Figure 8), since, in accordance with (24)

$$\omega L i(t) = \frac{U\omega t}{2}\sin \omega t. \tag{25}$$

But the projection of vector OA onto the X-axis defines only the first part of expression (25). The second part of (25) may be represented as the projection of vector AB = U/2, perpendicular to vector OA, so that the voltage $u_2(t)$ is expressed as the projection of the resulting vector OB and not of vector OA. It is evident that the single vector simultaneously defining quantities i(t) and $u_2(t)$ will be vector OC defined by the coordinates of $u_2(t)$ and i(t). It is easily shown that it is this vector which is defined by the resonance transform $D_\omega(t)$ and that point C belongs to a cycloid (Figure 6) rotating in common with direction OA given by the factor $e^{j\omega t}$.

6. DETERMINATION OF VARIABLE AMPLITUDE OF VOLTAGE AT CAPACITOR IN AN UNDAMPED CIRCUIT BY MEANS OF RESONANCE TRANSFORMS

The voltage at the capacitor (Figure 7) is equal to $u_1(t) = f(t) - u_2(t)$, where f(t) is the value of $f(\tau)$ (i.e., of the input voltage at moment t). Inserting the quantity $u_2(t)$ from expression (18) we obtain after transposition

$$U_1(t) = -j\omega F_\omega(t)\Big|_{\omega=\omega_K} = \frac{\omega}{j} F_\omega(t)\Big|_{\omega=\omega_K}, \tag{26}$$

whence we obtain

Theorem 9. Amplitude of the voltage at the capacitor is defined by the vector of the principal resonance transform multiplied by $-j\omega$ with $\omega = \omega_K$.

7. RELATION BETWEEN VARIABLE AMPLITUDES AND SPECTRAL DENSITY

Comparing expressions (13) and (20), we obtain

$$S_d(\omega) = D_\omega(t) = U_2(t)\,|_{\omega_K=\omega}, \tag{27}$$

that is, the instantaneous value of the current spectral density of the derivative for moment t_i is equal to the instantaneous value of the amplitude of voltage at the inductor as determined by the transform of $D_\omega(t_i)$ for the tuning frequency $\omega_K = \omega$.

Spectral density of the derivative for the completed process is determined by the amplitude of undamped voltage $U_2(\tau_k)$ in the circuit after completion of the process.

Comparing (14) and (26) we obtain

$$S_t(\omega) = F_\omega(t) = \frac{j}{\omega} U_1(t)\,|_{\omega_K=\omega}, \tag{28}$$

that is, the instantaneous value of spectral density is determined by the instantaneous value of amplitude of voltage at the capacitor multipled by j/ω. The spectral density of function $f(\tau)$ is determined by the amplitude of the undamped voltage $U_1(\tau_K)$ after completion of the process

$$S(\omega) = F_\omega(\tau_K) = \frac{j}{\omega} U_1(\tau_K)\,|_{\omega_K=\omega}. \tag{29}$$

8. RELATION BETWEEN THE LAPLACE TRANSFORM AND RESONANCE TRANSFORMS

The Laplace transform, given by the expression

$$f(p) = \int_0^\infty f(\tau)\, e^{-jp\tau} d\tau, \tag{30}$$

transforms the time function $f(\tau)$ to a function f(p). If in f(p) we insert the concrete value $p = \alpha + j\omega$, then we obtain the corresponding concrete value of f(p). This same value of f(p) is obtained by inserting in the integral of (30) the given value of $p = \alpha + j\omega$.

$$f(p) = \int_0^\infty f(\tau)\, e^{-\alpha\tau} e^{-j\omega\tau} d\tau \tag{31}$$

and integrating the function $f(\tau)\, e^{-\alpha\tau} d\tau$ in directions $-\omega\tau$ and thus solving the integral

$$\int_0^t f(\tau)\, e^{-\alpha\tau}\, e^{-j\omega\tau}\, d\tau$$

with a constant increase in t. In other words, the concrete value of f)p) may be regarded as the limit of the resonance transform of function $f(\tau)e^{-\alpha\tau}$ when $t \to \infty$.

The following theorem is easily proved.

Theorem 10. The value of the Laplace transform of a function for a given $p = \alpha + j\omega$ is equal to the limit (multiplied by j/ω) approached, for $t \to \infty$, by the amplitude of voltage at the capacitor in a lossless circuit to the input of which there is applied a voltage $u = f(\tau)e^{-\alpha\tau}$ (or is equal to the limit of the amplitude of voltage at the inductor divided by $j\omega$).

9. DAMPED RESONANCE TRANSFORMS

As has been shown in the previous sections, undamped resonance transforms are the solution of transients in a lossless circuit with voltage $u = f(\tau)$ applied to the input. The method of solving transients by means of resonance transforms may be extended to damped circuits [1]. The method of the solution (Figure 9) consists in subtracting from the values of input voltage $u = f(\tau)$ the values of voltage drop across an active resistance. Then we may consider that there acts on an undamped circuit an input voltage $u_3(t) = f(t) - i(t)r$ and the equation takes the form

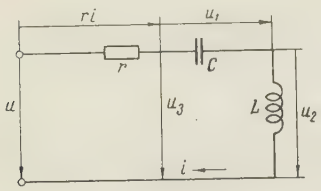


Figure 9. Circuit with active resistance.

$$L\frac{di}{df} + \frac{1}{C}\int_0^t i\,df = f(t) - i(t)\,r. \qquad (32)$$

consequently, the damped resonance transforms may be represented as

$$D_{\omega r}(t) = \int_0^t (f(\tau) - ir)'\, e^{-j\omega\tau}\, d\tau, \qquad (33)$$

$$F_{\omega r}(t) = \int_0^t (f(\tau) - ir)\, e^{-j\omega\tau}\, d\tau. \qquad (34)$$

The images of $D_{\omega r}(t)$ may easily be obtained by means of the device described in reference [1] Resistor r may be set over a wide range of positive or negative values by continuous adjustment of the mechanical feedback factor in the device. The device may be improved so that the feedback factor is automatically varied with a change in current or in voltage at the inductor. It will then be possible to define certain nonlinear curcuits on such a device.

10. SOME APPLICATIONS OF THE RESONANCE TRANSFORM METHOD

Let us comment on some applications of resonance transforms. As has already been stated, undamped resonance transforms are most closely related to the values of spectral density of the investigated process $f(\tau)$. Hence, in some certain cases the method of resonance transforms is extremely convenient for determining spectral functions and the currer spectrum. Thus, if $f(\tau)$ consists of any sequence of straight-line segments, then the resonanc transforms and, consequently, the values of the spectral function may be plotted by means of a compass and simple calculations.

In the case where the process $f(\tau)$ is given graphically (e.g. ,as the result of measurement), determination of the spectral function is considerably simplified in using the device cited in reference [1] in which mechanization of transformation of (1) is achieved kinematically. In addition to this instrument, which was demonstrated in improved form at the fair in Brussels, the principles of other kinematic instruments have been developed for mechanization of transformations of (1) and (2), permitting the solution of direct and inverse Fourier transforms. Moreover, on the basis of the theorems of constancy of image length and the radius of curvature, the author has developed a group of devices employing the analogy of flexible sheets (reference [4]) the width of which may be determined (by calculation or by

instrument) so that their line of curvature will correspond to the resonance image.

However, the role of resonance transforms is not limited to assisting in the solution of transient processes by means of the classical Fourier method. Undamped resonance transforms give the direct solution of a transient in a lossless circuit; that is, by use of the instrument assembly cited in reference [1], they permit solution of equations of the form

$$L\frac{di}{dt} + \frac{1}{C}\int idt = f(t)$$

with any right-hand member given graphically. This same instrument may be used to obtain damped resonance transforms, that is, to define the transient in a lossy circuit and to solve equations of the type

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int idt = f(t).$$

By means of damped resonance transforms we may also define the transients in more complex circuits when the transfer characteristic of the circuit may be presented in the form of a sum of damped oscillations. Then the resulting transient will appear as the sum of the transients in several isolated, damped circuits to the input of which there is applied the given process $f(\tau)$. These isolated transients are easily obtained on the instrument described in [1].

As has already been stated, the method of damped oscillations may also be extended to the solution of certain parametric and nonlinear problems. It is also noted that the method of resonance transforms is finding further applications.

REFERENCES

1. S. Dyad'kov, Resonance analogy, properties and mechanization of calculation of Fourier transforms, Acta technica, 1956, No. 6, Nakl. CSAV, Prague.
2. S. Djadkov, Spektralni funkce a resonancni krivky neperiodickych jevu, Sbornik praci fakulty electrotechnickeho inzenyrstvi za rok 1956, SNTL, Prague
3. I. M. Rabinovich, Obshchaya prochnost' i ustoychivost' sooruzheniy pri deystvii vzryvnoy nagruzki [Overall strength and stability of structures under rupture loads]; Symposium, No. 1, Gosudarstvennoye izdatel'stvo stroitel'noy literatury, 1944.
4. S. Djadkov, Resonanz-Transformationen, ihre Analogien und Anwendungen, IV. Internationales Kolloquium an der Hochschule fur Elektrotechnik Ilmenau (GDR), 1959.
5. Ya. Z. Tsypkin, Teoriya releynykh sistem avtomaticheskogo regulirovaniya [Analysis of relay systems for automatic control]: GITTL, 1955.

Submitted to the editors 13 May 1960

APPROXIMATE CALCULATION OF EIGENVALUES OF HIGHER MODES IN STRIP TRANSMISSION LINES

G. P. Samuilov

The report discusses a method of solving the problem of strip transmission lines which is based on reduction of the longitudinal propagation of the wave to transverse propagation with subsequent use of the method of functional equations. Design curves and experimental curves of phase velocity are presented.

INTRODUCTION

Strip transmission lines with the H_{10} mode are now under development, for the symmetrical strip lines with dielectric filler used in practice have high losses (Figure 1a). The phase velocity in such a system possesses dispersion. For calculation of the phase constant h let us use the relationship

$$\overline{h} = k\sqrt{1-(k_{кр}/k)^2} \qquad (k = 2\pi/\lambda). \tag{1}$$

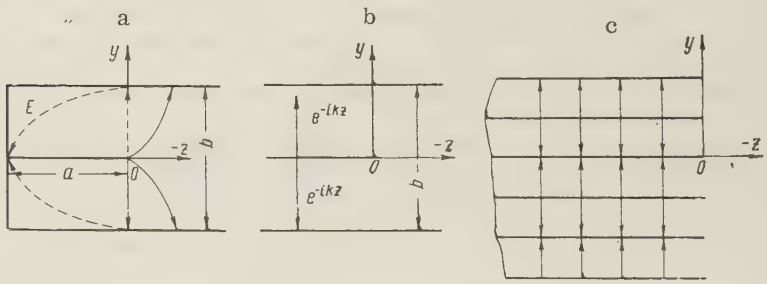


Figure 1. a) Strip line with H_{10} mode; b, c,) for calculation of phase velocity in strip line with H_{10} wave.

In formula (1) the critical wave number k_{cr} is unknown. For its calculation it is convenient to use a method permitting reduction of the tridimensional problem of longitudinal propagation of the wave to the bidimensional problem of its transverse propagation. A brief description of the method follows. Let h = 0. There will be no propagation in the longitudinal direction and in the transverse section of the line a standing wave will be established (Figure 1a). This wave may be represented as the result of addition of two antiphasal waves of equal amplitude with wave number k_{cr} and incident at the edge of the central half-plane (Figure 1b). In addition, in space ($z < 0$) the component fields E_y and H_x are attenuated due to their opposite orientation and this attenuation occurs in the direction of $z < 0$. At the shielding planes $y = \pm b/2$ the E_z component must satisfy the zero boundary conditions. In addition, if $b < \lambda/2$, the wave with such components will be attenuated and, the smaller the

ratio b/a, the greater this attenuation.

On the basis of the above observations we may conclude that radiation in the transverse direction (in direction of $z < 0$) is absent. In virtue of this the modulus of the reflection coefficient of the wave reflected from the edge of the central half-plane is equal to unity for each "partial" wave (above the half-plane and below it). Thus, a pure standing wave is established above and below the half-plane.

In this case, at a certain distance in the plane $z = a$ there must occur the "first null" of the electric field. If in the plane $z = a$ we place a metal partition, we shall obtain an entirely metallic strip line in which the H_{10} mode is propagated (Figure 1a). In the region of $z > 0$ the electric field may be represented in the form

$$E_y(z) = A\{e_{\text{}}^{-ik_{\text{кр}}z} + e^{ik_{\text{кр}}z+i\varphi_{00}}\}, \tag{2}$$

where φ_{00} is the phase of the reflection coefficient of the wave reflected from the edge of the central half-plane and depends on parameter b/a.

In accordance with the definition of point $z = a$, we may write

$$E_y(a) = 2Ae^{i\varphi_{00}/2}\cos\left(k_{\text{cr}}a + \frac{\varphi_{00}}{2}\right) = 0, \tag{3}$$

whence

$$\frac{k_{\text{cr}}a}{2\pi} = \frac{a}{\lambda_{\text{кр}}} - \frac{\varphi_{00}}{4\pi}. \tag{4}$$

Thus, calculating the angle φ_{00} as the result of solution of the bidimensional problem of diffraction of antiphasal waves of equal amplitude with wave number k_{cr} at the edge of a metal half-plane located between infinite planes, we may find from formulas (4) and (1) the phase constant of the wave which is propagated in the longitudinal direction and corresponds to solution of the tridimensional problem of a strip line with H_{10} mode.

It may be shown that with a distance $b > \lambda/2$ the strip wave is not localized in the region $0 < z < a$ and is radiated in the transverse direction in the space between planes. Such lines are not suitable for transmission of microwave energy and it is henceforth assumed that $b < \ll \lambda/2$.

1. PRINCIPAL INTEGRAL EQUATION AND METHOD OF SOLUTION (CALCULATION OF PHASE ANGLE φ_{00})

At the edge of a half-plane let there be incident a wave with wave number $k = k_{cr}$ (hereafter simply designated as k) in which only two components E_y, H_x differ from zero.

In approaching the "open end" the propagating wave preserves its corresponding symmetry of current distribution; hence among the reflected waves we find waves of only this type.

Let us project the mirror image of relatively infinite planes (Figure 1c). The currents at the sheets satisfy the condition

$$I_n = -I_{n+1} = -I_{n-1}, \tag{5}$$

while the current distribution function remains constant from sheet to sheet.

As the result of simple transformation we obtain the integral equation

$$\left(\frac{\partial^2}{\partial z^2} + k^2\right)\int_0^\infty K(z-\zeta)f(\zeta)\,d\zeta = 0 \quad (z>0), \tag{6}$$

where

$$K(z) = \frac{1}{2}\sum_{n=-\infty}^{n=+\infty}(-1)^n H_0^{(1)}\left(k\sqrt{z^2+(nb)^2}\right) \tag{7}$$

For solution of (3) L. A. Vaynshteyn developed an extremely simple method (reference [1]) according to which the required function f is represented in the form of a sum:

$$f(z) = A\{e^{-ikz} + R_{00}e^{ikz} + \sum_{l\neq 0}^{\infty} R_{0l}\,e^{iw_l z}\},$$

where

$$R_{00} = -\frac{G_2(-k)}{G_2(k)} = -\frac{G_1(k)}{G_2(k)};$$

$$R_{0l} = -\frac{2kG_1(k)}{(w_l^2 - k^2)\, G_2'(k)}; \tag{8}$$

$$G(w) = \int_{-\infty}^{+\infty} e^{-iwz} K(z)\, dz\, (w^2 - k^2) =$$

$$= (w^2 - k^2) G_1(w) G_2(w);$$

$$G(w_l) = 0.$$

The complete expression for current in reference [1] also includes the integral for the vertical cross section.

In the given case it may be shown that this integral is equal to zero, since the currents characterized by it, flowing on opposite sides of the sheet, cancel one another. Functions $G_1(w)$ and $G_2(w)$ must be holomorphic in the upper and lower half-planes, respectively, and must have no zeroes there.

In order to find the function $G(w)$ by using the Poisson theorem let us transform (4) to the following form:

$$K(z) = \frac{bi}{2} \sum_{n=0}^{\infty} \frac{e^{-\frac{2\pi z}{b}\sqrt{\left(n+\frac{1}{2}\right)^2 - \left(\frac{kb}{2\pi}\right)^2}}}{\sqrt{\left(n+\frac{1}{2}\right)^2 - \left(\frac{kb}{2\pi}\right)^2}}. \tag{9}$$

In accordance with (8), the function $G_0(w) = G(w)/(w^2 - k^2)$ is represented by the series

$$G_0(w) = \frac{bi}{2} \sum_{n=0}^{\infty} \frac{1}{\left[(2n+1)\frac{\pi}{2}\right]^2 - \left(\frac{kb}{2}\right)^2 + \left(\frac{wb}{2}\right)^2}. \tag{10}$$

Summation of the latter is easily obtained by means of the Mittag-Loeffler formula:

$$G_0(w) = \frac{1}{v} \frac{1 - e^{ivb}}{1 + e^{ivb}}, \tag{11}$$

where

$$v = \sqrt{k^2 - w^2}.$$

Using the results of reference [1], let us define the functions $G_1(w)$ and $G_2(w)$:

$$G_{1,2}(w) = \sqrt{\frac{-ib \operatorname{tg} \pi q}{\pi q}} \sqrt{k \pm w} \frac{\prod_{m=1}^{\infty} \left(1 \pm \frac{w'}{\delta_m}\right) e^{(\pm\mu_m \mp \mu_{m-1})w'}}{\prod_{m=1}^{\infty} \left(1 \pm \frac{w'}{\gamma_m}\right) e^{(\pm\tau_m \mp \tau_{m-1})w'}}, \tag{12}$$

where

$$q = b/\lambda;$$

$$\gamma_m = \sqrt{q^2 - \left(m - \frac{1}{2}\right)^2} \quad (\operatorname{Im} \gamma_m > 0),\ m = 1; 2; 3; \ldots;$$

$$q \sin \tau_m = \gamma_m, \quad \tau_0 = \pi/2;$$

$$\delta_m = \sqrt{q^2 - m^2}, \quad m = 1; 2; 3; \ldots; \tag{13}$$

$$q \sin \mu_m = \delta_m, \quad \mu_0 = \pi/2; \quad w' = \frac{w\lambda}{2\pi}.$$

With analysis of (10) it is not difficult to conclude that the conditions which functions $G_{1,2}(w)$ must satisfy are fulfilled.

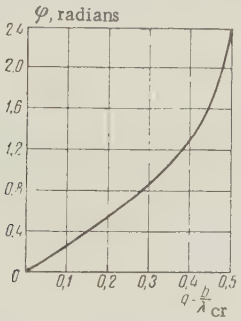
On the basis of formulas (8) and (12) we may find the phase of the reflection coefficient

$$\varphi_{00} = 2\sum_{m=1}^{\infty}\left\{\arcsin\frac{q}{m-\frac{1}{2}} - \arcsin\frac{q}{m}\right\}. \qquad (14)$$

It is not difficult to verify by direct check that the modulus of the reflection coefficient $|R_{00}|$ is equal to unity; this confirms the statements in the introduction.

2. A REMARK ON THE NATURE OF THE FIELD IN A STRIP LINE WITH H_{10} MODE AND EVALUATION OF NUMERICAL RESULTS

From the theory of symmetrical strip lines of conventional type it is known that in the region of the central conductor there is localized almost all the microwave energy transmitted along the strip line



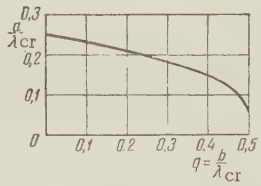


Рис. 3

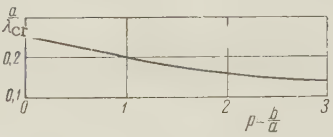


Figure 2. Phase of reflection coefficient φ_{00} as a function of $q = b/\lambda_{cr}$.
Figure 3. Critical wavelength as a function of $q = b/\lambda_{cr}$.
Figure 4. Critical wavelength as a function of $P = b/a$.

Transverse distribution of the field in this region is the sum of a uniform field and the field caused by the influence of the edges of the central conductor. The smaller this influence, the smaller the ratio b/a.

Beyond the region of the central conductor the transverse distribution of the field is purely attenuated and this attenuation increases as b/a decreases. In the case of a strip line with H_{10} mode we may also show that transverse distribution of the field in the region $0 < z < a$ consists of the field of a wave which is described by a formula which is essentially similar to formula (2) and of an attenuated field caused by the influence of the edge of the half-plane. With an increase in z, the smaller ratio b/a, the more rapidly the magnitude of the latter field decreases. The field in the region $z < 0$ is also damped and with an increase in $|z|$, the smaller b/a, the more rapidly this field decreases.

The approximation of our present method lies in neglecting the attenuated field and defining the plane $z = a$ of the zero field E_y only in terms of the standing wave of type (2). The associated error is evidently small in virtue of the rapid decay of the damped field.

Figure 2 represents the phase φ_{00} as a function of $q = b/\lambda_{cr}$. With $q \ll 1$ the phase of the reflection coefficient is equal to zero, which corresponds to the reflection from a "magnetic wall." In this case the field is almost entirely localized in the region of the central conductor. With an increase in q the phase also increases. This attests to the fact that the field issues from the region occupied by the "strip."

Figure 3 shows the dependency of ratio a/λ_{cr} on b/λ_{cr} for the H_{10} mode. It is evident that with small b/λ_{cr} the strip line behaves as a waveguide in which one of the walls

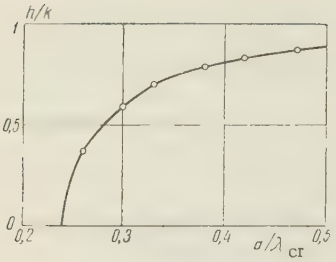


Figure 5. Relative phase constant in strip line with H_{10} mode and b/a = 0.28 (points indicate data of experiment).

is magnetic and the other metallic. A plot of a/λ_{cr} as a function of b/a (Figure 4) is required for practical calculations. By using formula 1 we may calculate the phase constant in a strip line at any frequency. The design values and experimental values for relative phase constant h/k, derived for the ratio b/a = 0.28 and presented here for purposes of illustration (Figure 5), show high coincidence.

In conclusion the author expresses his thanks to N.G. Trenev for consultation in organizing the presentation of this subject.

REFERENCES

1. L.A. Vaynshteyn, Difraktsiya elektromagnitnykh i zvukovykh voln na otkrytom kontse volnovoda, [Diffraction of electromagnetic and sound waves at the open end of a waveguide]; Izd. Sovyetskoye radio, 1950.

Submitted to the editors 23 May 1960

CHARACTERISTIC IMPEDANCE OF MULTIWIRE LINES WITH CIRCULAR CONDUCTORS

V. M. Dashenkov

This report presents formulas and tables for calculation of the characteristic impedanc of single-row, stacked row and bidimensional-periodic lines with conductors of circular cross section located along mutually perpendicular directions. Knowledge of this characteri stic impedance is necessary in investigating the dispersion and coupling impedance of s slow-wave lines. On the basis of electrostatic potential coefficients formulas are derived for the case where the distance between conductors is considerably greater than their radii. It is assumed that the formulas may be used if the ratio of the diameter of the conductors to the distance between them does not exceed 0.3.

INTRODUCTION

In the calculation of stub systems by the method of multiwire lines it is necessary to know the quantities referred to as the characteristic impedances of the multiwire line, whic are functions of the cross-sectional parameters of the conductors and the phase-shift angle between them.

For different multiwire lines with conductors of rectangular cross section the design formulas for characteristic impedances are given in references [1-5]. For lines with conductors of circular cross section appropriate results are known only for single-row systems without side walls from references [6] (sufficiently exact solution) and [7] (approximate solution). However, the formulas in reference [6] are somewhat complex for practical use.

For conductors of circular cross section the present report presents extremely

simple formulas for the calculation of characteristic impedances of bidimensional-periodic (Figure 1), stacked (Figures 2, 4) and single-row multiwire lines with side wall and without it (Figure 3).

1. GENERAL RELATIONSHIPS

Figure 1 represents the cross section of a multirow bidimensional-periodic) line with periods D_z and D_y and conductor radius r. Each conductor is assigned a two-digit number pq which indicates that it belongs to the p-th horizontal and the q-th vertical rows.

If the number of rows in the y direction is limited (e.g., p = 1, 2, ..., Q = 1), the system is referred to as a Q-row system. In this case, as shown in reference [8], it is described by a square matrix of characteristic impedances of the Q-th order

$$\tilde{K} = \left\| K_{ik}(\psi) \right\|_{0}^{Q-1},$$

where

$$K_{ii} = \frac{1}{c} \sum_{q=-\infty}^{\infty} \alpha_{isiq} e^{-j\psi(q-s)}; \tag{1}$$

$$K_{ik} = \frac{1}{c} \sum_{q=-\infty}^{\infty} \alpha_{iskq} e^{-j\psi(q-s)} \tag{2}$$

are quantities which may be referred to respectively as the natural characteristic impedance of the i-th row and the mutual characteristic impedance of the i-th and k-th rows. In these formulas c is the free-space velocity of light; ψ is the phase difference of oscillations between adjacent conductors in the z direction; α_{iskg} and α_{isig} are the potential coefficients of the line, wherein s and q are integers characterizing the number of the vertical row. The characteristic impedances are independent of s; hence s may have any value, e.g., s = 0

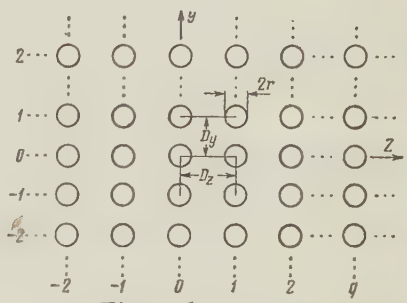


Figure 1

Special cases of the stacked line without walls are the three-row and two-row structures (Figure 2). The number of possible wave modes propagated in a stacked line is equal to the number of rows. For each wave mode we may indicate the corresponding characteristic impedance $K_n(\psi)$ (n is the wave number), which is the root of the characteristic equation of the $\tilde{K}$ matrix.

It is shown in reference [8] that the characteristic impedances of a three-row line (Figure 2a) will be

$$K_1(\psi) = K_{00}(\psi) + \frac{1}{2} K_{02}(\psi) + \sqrt{2K_{01}^2(\psi) + \frac{1}{4} K_{02}^2(\psi)} \qquad \text{(synphasic)} \tag{3}$$

$$K_2(\psi) = K_{00}(\psi) - K_{02}(\psi), \tag{4}$$

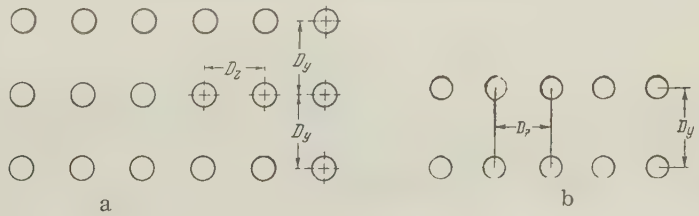


Figure 2.

$$K_3(\psi) = K_{00}(\psi) + \frac{1}{2} K_{02}(\psi) - \sqrt{2K_{01}^2(\psi) + \frac{1}{4} K_{02}^2(\psi)} \qquad \text{(antiphasic)} \tag{5}$$

For a two-row line (Figure 2b) we obtain

$$K_1(\psi) = K_{00}(\psi) + K_{01}(\psi) \qquad \text{(synphasic)} \tag{6}$$

$$K_2(\psi) = K_{00}(\psi) - K_{01}(\psi) \qquad \text{(antiphasic)} \tag{7}$$

The last formula also describes the characteristic impedance of a single-row system with side wall (Figure 3).

The characteristic impedance of a bidimensional-periodic line (with number of rows $Q = \infty$) is defined by the formula (references [8, 9])

$$K(\psi, \theta) = \frac{1}{c} \sum_{q,\, k=-\infty}^{\infty} \alpha_{iskq}\, e^{-j[\psi(q-s)+\theta(k-i)]}, \tag{8}$$

where θ is the phase shift of the wave between adjacent rows in the y direction. With consideration of (1) and (2)

$$K(\psi, \theta) = \sum_{k=-\infty}^{\infty} K_{ik}(\psi)\, e^{-j\theta(k-i)}. \tag{9}$$

The characteristic impedances of the various modes of a Q-row line bounded by metal walls (y=const) which are located at distances of $D_y/2$ from the extreme rows (Figure 4), as shown in reference [9], are equal to $K(\psi, \theta)$ with $\theta = n\pi/\theta$, where $n = 1, 2, \ldots, Q$.

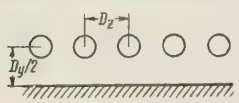


Figure 3

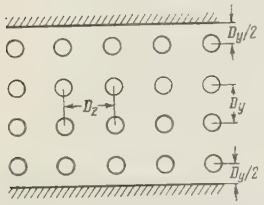


Figure 4

Due to the fact that between the potential coefficients of a stacked line the following relationships apply

$$\alpha_{iskq} = \alpha_{iqks} = \alpha_{ksiq} = \alpha_{kqis}, \tag{10}$$

the characteristic impedances of (1) and (2) are even functions and satisfy the reciprocity theorem:

$$K_{ik}(\psi) = K_{ik}(-\psi), \quad K_{ik}(\psi) = K_{ki}(\psi). \tag{11}$$

This, in turn, indicates the $K_{ik}(\psi)$ are real values. For a bidimensional-periodic line, as the complement of (10) by virtue of the identity of all rows,

$$\alpha_{iskq} = \alpha_{i+1,\, sk+1,\, p}\,, \alpha_{0\,skq} = \alpha_{0s,-kq} \tag{12}$$

and, hence,

$$K_{ik}(\psi) = K_{0p}(\psi) = K_{0,-p}(\psi) \quad (p = k-i). \tag{13}$$

Keeping (11) and (13) in mind, formula (9) may be presented in the form

$$K(\psi, \theta) = K_{00}(\psi) + 2 \sum_{p=1}^{\infty} K_{0p}(\psi) \cos p\theta, \tag{14}$$

wherein $K_{00}(\psi)$ and $K_{0p}(\psi)$, taking into account (10) and (12), may be expressed in similar form:

$$K_{00}(\psi) = \frac{1}{c} (\alpha_{0000} + 2 \sum_{q=1}^{\infty} \alpha_{000\,q} \cos q\psi), \tag{15}$$

$$K_{0p}(\psi) = \frac{1}{c} (\alpha_{00\,p0} + 2 \sum_{q=1}^{\infty} \alpha_{00\,pq} \cos q\psi). \tag{16}$$

We will note that as long as the values of the potential coefficients are correctly defined, formulas (14)-(16) are rigorous. Thus, the choice in as convenient a form as possible, of one or another approximation for the value of the potential coefficients is of greatest importance in the calculation of the characteristic impedances.

We shall begin calculation of $K_{00}(\psi)$ and $K_{0p}(\psi)$ with their transformation to the following form:

$$K_{00}(\psi) = \frac{1}{2\sin^2\frac{\psi}{2}\,c}\{(\alpha_{0000} - \alpha_{0001}) - \sum_{q=1}^{\infty}[(\alpha_{000,q-1} - \alpha_{000q}) - (\alpha_{000\,q} - \alpha_{000,\,q+1})]\cos q\psi\}, \tag{17}$$

$$K_{0p}(\psi) = \frac{1}{2\sin^2\frac{\psi}{2}\,c}\{(\alpha_{00\,p0} - \alpha_{00\,p1}) - \sum_{q=1}^{\infty}[(\alpha_{00\,p,\,q-1} - \alpha_{00\,pq}) - (\alpha_{00\,pq} - \alpha_{00\,p,\,q+1})]\cos q\psi\}. \tag{18}$$

According to the definition of potential coefficients in reference [10] it is not difficult to see that $\frac{1}{c}(\alpha_{0000} - \alpha_{00\,pq})$ is half of the characteristic impedance of a two-wire line formed by conductors 00 and pq (it is assumed that the radii of the conductors are identical). It is evident that the effect of the presence of all the remaining conductors on the value of this impedance cannot be taken into account if D_z, $D_y \gg r$ (reference [10]). This case is also discussed below.

Then

$$\frac{1}{c}(\alpha_{0000} - \alpha_{00\,pq}) = 60\ln\frac{\sqrt{(qD_z)^2 + (pD_y)^2}}{r} = 60\ln\frac{D_z}{r}\sqrt{q^2 + p_1^2}\ \text{ohms}, \tag{19}$$

where

$$p_1 = \delta p; \qquad \delta = \frac{D_y}{D_z}. \tag{20}$$

Similarly

$$\frac{1}{c}(\alpha_{00\,pq} - \alpha_{00\,p,q+1}) = \frac{1}{c}(\alpha_{0000} - \alpha_{00\,p,\,q+1}) - \frac{1}{c}(\alpha_{0000} - \alpha_{00\,pq}) =$$

$$= 60\ln\frac{D_z}{r}\sqrt{p_1^2 + (q+1)^2} - 60\ln\frac{D_z}{r}\sqrt{p_1^2 + q^2} = 30\ln\frac{p_1^2 + (q+1)^2}{p_1^2 + q^2}\ \text{ohms} \tag{21}$$

$$\frac{1}{c}(\alpha_{00\,p,\,q-1} - \alpha_{00\,pq}) = 30\ln\frac{p_1^2 + q^2}{p_1^2 + (q-1)^2}\ \text{ohms}, \tag{22}$$

$$\frac{1}{c}(\alpha_{000\,q} - \alpha_{000,\,q+1}) = 60\ln\frac{q+1}{q}\ \text{ohms}, \tag{23}$$

$$\frac{1}{c}(\alpha_{000,\,q-1} - \alpha_{000\,q}) = 60\ln\frac{q}{q-1}\ \text{ohms}, \tag{24}$$

$$\frac{1}{c}(\alpha_{0000} - \alpha_{0001}) = 60\ln\frac{D_z}{r}\ \text{ohms}. \tag{25}$$

Instead of (22), we may use the formula

$$\frac{1}{c}(\alpha_{0000} - \alpha_{0001}) = 60\,\mathrm{Arch}\frac{D_z}{2r}\ \text{ohms}, \tag{26}$$

which, as is known, is valid for any D_z/r. In this formula we consider the "proximity effect" between adjacent conductors of the row. In view of the above observations, in the given approximation [i.e., in using (20)-(26)] the value of $K_{00}(\psi)$ is equal to the impedance of one isolated row and $K_{0p}(\psi)$ is equal to the mutual impedance of isolated rows O and P.

2. CALCULATION OF IMPEDANCE $K_{00}(\psi)$

Inserting (23,) (24) and (26) into (17) we obtain

$$K_{00}(\psi) = \frac{60}{2\sin^2\frac{\psi}{2}}\left(2\sin^2\frac{\psi}{2}\,\mathrm{Arch}\frac{D_r}{2r} + \cos\psi\ln 2 - \sum_{q=2}^{\infty}\ln\frac{q^2}{q^2-1}\cos q\psi\right)\ \text{ohms}. \tag{27}$$

Without dwelling on the intermediate calculations, let us present $K_{00}(\psi)$ in the following form, suitable for practical calculations:

$$K_{00}(\psi) = 60\left[\operatorname{Arch}\frac{D_z}{2r} - 1 + S(\psi)\right] \text{ ом}, \tag{28}$$

where

$$S(\psi) = \frac{1}{2\sin^2\frac{\psi}{2}}\left[(1+\ln 2)\cos\psi + \frac{\pi}{\sqrt{2}}\frac{\cos\frac{\pi-\psi}{\sqrt{2}}}{\sin\frac{\pi}{\sqrt{2}}} - \right.$$

$$\left. - \sum_{q=2}^{\infty}\left(\ln\frac{q^2}{q^2-1} - \frac{1}{q^2-1}\right)\cos q\psi\right] \quad (0 \leqslant \psi < \pi). \tag{29}$$

In distinction from the series in (27), the series in (29) converges rapidly and the relative weight of its terms is not great. Function $S(\psi)$ is easily calculated. Its values are given in Table 1.

Table 1

ψ/π	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$S(\psi)$	∞	9,165	4,174	2,523	1,712	1,241	0,9461	0,7556	0,6358	0,5691	0,5484

Comparison of formula (28) with the corresponding data of exact calculation from refere[6] shows that with $\psi > 0.7\pi$ the formula yields higher values and with $\psi > 0.7\pi$ lower values of $K_{00}(\psi)$. The maximum error observed with $\psi \simeq 0.3$ increases with an increase in $2r/D_z$ and at $2r/D_z = 0.1$, 0.3, 0.5, and 0.7 is 0.3, 5, 10.5, and 33% respectively.

3. CALCULATION OF IMPEDANCE $K_{0p}(\psi)$

Insertion of (20)-(22) into (18) yields

$$K_{0p}(\psi) = \frac{30}{2\sin^2\frac{\psi}{2}}\left(\ln\frac{1+p_1^2}{p_1^2} + \sum_{q=1}^{\infty}\ln\frac{[p_1^2+(q-1)^2][p_1^2+(q+1)^2]}{(p_1^2+q^2)^2}\cos q\psi\right) \text{ ohms}, \tag{30}$$

The given series converges slowly. However, if we differentiate it with respect to p_1 and then calculate the derived sums and integrate the result, it is then replaced by another extremely rapidly converging series.

$$K_{0p}(\psi) = 60\pi\sum_{q=0}^{\infty}\left(\frac{e^{-p_1(2\pi q+\psi)}}{2\pi q+\psi} + \frac{e^{-p_1[2\pi(q+1)-\psi]}}{2\pi(q+1)-\psi}\right) \quad (0 < \psi \leqslant \pi), \tag{31}$$

in which the retention of only two terms (q= 0 and 1) with $p_1 = \delta p = \frac{D_y}{D_z}p \geqslant 0.2$ results in an error of less than 2% in comparison with the exact value of $K_{0p}(\psi)$. Three terms result in ar accuracy of not less than 3% for $p_1 \geqslant 0.1$.

The values of quantity $1/60K_{0p}(\psi)$ calculated by means of formula (31) are listed in Tabl

4. CALCULATION OF IMPEDANCE $K(\psi, \theta)$

From (28), formula (14(is conveniently presented in the following form:

$$K(\psi, \theta) = 60\left[\operatorname{Arch}\frac{D_z}{2r} - 1 + S(\psi) + S(\delta, \psi, \theta)\right], \tag{32}$$

where

$$S(\delta, \psi, \theta) = \frac{2}{60}\sum_{p=1}^{\infty}K_{0p}(\psi)\cos p\theta. \tag{33}$$

Keeping in mind (31) and the formula in reference [11]

$$\sum_{k=1}^{\infty}e^{-kt}\cos kx = \frac{1}{2}\left(\frac{\operatorname{sh} t}{\operatorname{ch} t - \cos x} - 1\right) \quad (t > 0), \tag{34}$$

Table 2

δ_p	ψ/π										
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	
0,2	9,721	4,749	3,114	2,314	1,859	1,568	1,384	1,268	1,206	1,190	
0,3	9,267	4,314	2,695	1,911	1,463	1,213	1,011	0,9018	0,8416	0,8228	
0,4	8,904	3,979	2,384	1,621	1,190	0,9268	0,7612	0,6590	0,6032	0,5858	
0,5	8,592	3,702	2,135	1,395	0,9834	0,7368	0,5829	0,5286	0,4375	0,4219	
0,6	8,307	3,456	1,924	1,212	0,8229	0,5925	0,4512	0,3662	0,3205	0,3060	
0,7	8,027	3,235	1,741	1,059	0,6942	0,4800	0,3519	0,2759	0,2355	0,2229	
0,8	7,785	3,029	1,579	0,9272	0,5858	0,3909	0,2759	0,2088	0,1735	0,1621	
0,9	7,539	2,845	1,434	0,8143	0,4955	0,3195	0,2172	0,1585	0,1280	0,1186	
1,0	7,307	2,670	1,302	0,7160	0,4218	0,2620	0,1716	0,1206	0,09458	0,08635	
2,0	5,335	1,423	0,5064	0,2027	0,08635	0,03858	0,01781	$8,657 \cdot 10^{-3}$	$4,802 \cdot 10^{-3}$	$3,684 \cdot 10^{-3}$	
3,0	3,897	0,7597	0,1973	0,05769	0,01796	$5,842 \cdot 10^{-3}$	$1,952 \cdot 10^{-3}$	$6,760 \cdot 10^{-4}$	$2,593 \cdot 10^{-4}$	$1,618 \cdot 10^{-4}$	
4,0	2,847	0,4053	0,07693	0,01643	$3,684 \cdot 10^{-3}$	$8,876 \cdot 10^{-4}$	$2,167 \cdot 10^{-4}$	$5,420 \cdot 10^{-5}$	$1,457 \cdot 10^{-5}$	$6,998 \cdot 10^{-6}$	
5,0	2,080	0,2163	0,02998	$4,676 \cdot 10^{-3}$	$7,780 \cdot 10^{-4}$	$1,348 \cdot 10^{-4}$	$2,403 \cdot 10^{-5}$	$4,429 \cdot 10^{-6}$	$8,371 \cdot 10^{-7}$	$3,027 \cdot 10^{-7}$	
6,0	1,519	0,1154	0,01168	$1,330 \cdot 10^{-3}$	$1,618 \cdot 10^{-4}$	$2,048 \cdot 10^{-5}$	$2,668 \cdot 10^{-6}$	$3,547 \cdot 10^{-7}$	$4,877 \cdot 10^{-8}$	$1,309 \cdot 10^{-8}$	
7,0	1,110	0,06158	$4,555 \cdot 10^{-3}$	$3,790 \cdot 10^{-4}$	$3,365 \cdot 10^{-5}$	$3,112 \cdot 10^{-6}$	$2,900 \cdot 10^{-7}$	$2,874 \cdot 10^{-8}$	$2,860 \cdot 10^{-9}$	$5,636 \cdot 10^{-10}$	
8,0	0,8106	0,03286	$1,755 \cdot 10^{-3}$	$1,079 \cdot 10^{-4}$	$6,998 \cdot 10^{-6}$	$4,728 \cdot 10^{-8}$	$3,284 \cdot 10^{-8}$	$2,330 \cdot 10^{-9}$	$1,688 \cdot 10^{-10}$	$2,449 \cdot 10^{-11}$	
9,0	0,5921	0,01753	$6,920 \cdot 10^{-4}$	$3,073 \cdot 10^{-5}$	$1,455 \cdot 10^{-6}$	$7,181 \cdot 10^{-8}$	$3,645 \cdot 10^{-9}$	$1,888 \cdot 10^{-10}$	$9,967 \cdot 10^{-12}$	$1,059 \cdot 10^{-12}$	
10,0	0,4325	$9,354 \cdot 10^{-3}$	$2,697 \cdot 10^{-4}$	$8,748 \cdot 10^{-6}$	$3,027 \cdot 10^{-7}$	$1,091 \cdot 10^{-8}$	$4,045 \cdot 10^{-10}$	$1,531 \cdot 10^{-11}$	$5,893 \cdot 10^{-13}$	$4,581 \cdot 10^{-14}$	

Note: Comma designates decimal point.

we find

$$S(\delta, \psi, \theta) = \pi \sum_{q=0}^{\infty} \left[\frac{1}{2\pi q + \psi} \left(\frac{\operatorname{sh} \delta (2\pi q + \psi)}{\operatorname{ch} \delta (2\pi q + \psi) - \cos \theta} - 1 \right) + \right. \tag{34}$$

$$\left. + \frac{1}{2\pi(q+1) - \psi} \left(\frac{\operatorname{sh} \delta [2\pi (q+1) - \psi]}{\operatorname{ch} \delta [2\pi (q+1) - \psi] - \cos \theta} - 1 \right) \right] \quad (0 < \psi \leqslant \pi).$$

Table 3

θ/π	ψ/π										
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
0	∞	63,4	15,6	6,78	3,71	2,30	1,56	1,14	0,895	0,768	0,729
0,1	63,4	31,5	12,4	6,08	3,48	2,21	1,51	1,11	0,880	0,757	0,719
0,2	15,6	12,4	7,66	4,62	2,92	1,96	1,39	1,04	0,836	0,725	0,690
0,3	6,78	6,08	4,62	3,27	2,29	1,64	1,22	0,942	0,771	0,677	0,647
0,4	3,71	3,48	2,92	2,29	1,75	1,34	1,04	0,830	0,695	0,620	0,596
0,5	2,30	2,21	1,96	1,64	1,34	1,08	0,871	0,721	0,619	0,560	0,541
0,6	1,56	1,51	1,39	1,22	1,04	0,871	0,732	0,624	0,548	0,504	0,489
0,7	1,14	1,11	1,04	0,942	0,830	0,721	0,624	0,546	0,490	0,456	0,444
0,8	0,895	0,880	0,836	0,771	0,695	0,619	0,548	0,490	0,447	0,420	0,411
0,9	0,768	0,757	0,725	0,677	0,620	0,560	0,504	0,456	0,420	0,397	0,389
1,0	0,729	0,719	0,690	0,647	0,596	0,541	0,489	0,444	0,411	0,389	0,383

Note: Comma designates decimal point.

This series converges quite rapidly and for practical calculations in most cases it suffices to consider its first two terms (q = 0 and 1).

The numerical values of the sum of function $S(\psi) + S(\delta, \psi, \theta)$ for the case $\delta = D_y/D_z = 1$ are listed in Table 3.

5. ASYMPTOTIC VALUES OF CHARACTERISTIC IMPEDANCES WITH $\psi \ll \pi$, $\psi = 0$, $\psi = \pi$

Without dwelling on mathematical details, we shall point out that with $\psi \ll \pi$

$$K_{00}(\psi) = 60 \left(\operatorname{Arch} \frac{D_z}{2r} - \ln 2\pi + \frac{\pi}{\psi} \right) \simeq 60 \frac{\pi}{\psi}. \tag{35}$$

Physically, the given result is completely justified, for it is known from reference [12] that at small ψ any real single-row line behaves as an anisotropic conducting plane the characteristic impedance of which is also defined (reference [35]).

With $\psi = \pi$

$$K_{00}(\pi) = 60 \left(\operatorname{Arch} \frac{D_z}{2r} - \ln \frac{\pi}{2} \right).$$

Using formula (25) instead of (26), we obtain

$$K_{00}(\pi) = 60 \ln \frac{2D_z}{\pi r}, \tag{36}$$

which coincides with the known expression for the characteristic impedance of a line consisting of a conductor located midway between two parallel metal planes separated by a distance D_z (reference [13]).

From (28), at small values of ψ, we obtain

$$K_{0p}(\psi) = 60 \left(\frac{\pi}{\psi} - \ln 2 \operatorname{sh} p_1 \pi \right) \simeq 60 \pi \left(\frac{1}{\psi} - p_1 \right). \tag{37}$$

Hence, it is quite evident that at small values of ψ we cannot neglect the interaction between remote rows, since with an increase in p the impedance $K_{0p}(\psi)$ decreases slowly (see also Table 2). At the same time

$$K_{0p}(\pi) = 60 \ln \operatorname{cth} \frac{p_1 \pi}{2}, \tag{38}$$

whence it follows that with $\psi = \pi$ with an increase in p the interaction of the rows decreases rapidly and, consequently, in this case the influence of the remote rows may be neglected in calculation (see Table 2).

With $\psi \ll \pi$

$$S(\delta, \psi, \theta) = 60\left[-\frac{\pi}{\psi} + \frac{\delta\pi}{2\sin^2\frac{\theta}{2} + \frac{(\delta\psi)^2}{2}} + \sum_{q=1}^{\infty}\frac{1}{q}\left(\frac{\operatorname{sh} 2\pi\delta q}{\operatorname{ch} 2\pi\delta q - \cos\theta} - 1\right)\right]. \quad (39)$$

Inserting (35) and (39) into (32) and assuming that $\psi = 0$, we find

$$K(0, \theta) = 60\left[\operatorname{Arch}\frac{D_z}{2r} - \ln 2\pi + \frac{\delta\pi}{2\sin^2\frac{\theta}{2}} + \sum_{q=1}^{\infty}\frac{1}{q}\left(\frac{\operatorname{sh} 2\pi\delta q}{\operatorname{ch} 2\pi\delta q - \cos\theta} - 1\right)\right]. \quad (40)$$

The characteristic impedance of a single-row line with side wall (Figure 3) for $\psi \ll \pi$ and $\psi = 0$ is found from (7) from (35) and (37) and with use of (25)

$$K(0) = 60\ln\frac{D_z \operatorname{sh}\delta\pi}{r\pi}. \quad (41)$$

On the basis of (7), (36) and (38)

$$K(\pi) = 60\ln\left(\frac{2D_z}{\pi r}\operatorname{th}\frac{\delta\pi}{2}\right). \quad (42)$$

By means of (41) and (42) we may solve the problem of synthesis of the given line, that is, we may define its geometry from the given K(0) and K(π):

$$\left.\begin{aligned} \frac{D_z}{r} &= \frac{\pi}{2}\frac{b}{\sqrt{a^2-1}}, \\ \frac{D_y}{D_z} &= \delta = \frac{2}{\pi}\operatorname{Arch} a, \end{aligned}\right\} \quad (43)$$

where

$$a = \exp\frac{K(0) - K(\pi)}{120}; \quad b = \exp\frac{K(0) + K(\pi)}{120}$$

CONCLUSION

Absence in the literature of appropriate formulas does not permit evaluating the limiting geometrical dimensions of a system for which the formulas for $K_{00}(\psi)$ and $K(\psi, \theta)$ may still be used. However, since their derivation and the derivation of $K_{00}(\psi)$ is performed on identical assumptions, it is probable that the limits of their applicability are identical. Hence, it appears that the error in the formulas must not be excessive (not more than 5%) for $2r/D_z$, $2r/D_y \leqslant 0.3$. For D_z, $D_y \gg r$ all the formulas are accurate.

The author is indebted to V.N. Ivanov for indicating the possibility of improving the convergence of the series (27) and to all participants in the radio physics seminar of the Saratov State University who joined in evaluation of this work.

REFERENCES

1. R.C. Fletcher, Proc, IRE, 1952, 40, 8, 951.
2. J.C. Walling, J. Electronics and Control, 1957, 3, 3, 239.
3. I.Sh. Beluga, Izv. vuzov MVO SSSR (Radiofizika), 1959, 2, 1, 84.
4. P.N. Butcher, Proc, IRE, part B, 1957, 104, 14, 177,
5. V.N. Ivanov, Radiotekhnika i elektronika, 1960, 5, 2, 224.
6. Chiao-Min-Chy, J. Appl. Phys, 1958, 29, 1, 88.
7. V.M. Dashenkov, Toward a theory of bidimensional stub slow-wave systems, Dokl. na gorodsk. nauchn. konferentsii po radioelektonika [Reports at the Municipal Scientific Conference on Radio Electronics], Saratov, 15 October 1959.
8. V.M. Dashenkov, Izv. vuzov MVO SSSR (Radiofizika) 1958, 1, 5-6, 121.
9. R.A. Silin, (Collection) Sb., Elektronika, 1958, 2, 3.

10. K.A. Krug, Osnovy elektrotekhniki, [Principles of Electrical Engineering]; GEI, 1952. 316.
11. I.M. Ryzhik, I.S. Gradshteyn, Tablitsy integralov, summ, ryadov i proizvedeniy, [Tables of integrals, sums, series and products]; GITTL, 1951.
12. Traveling-Wave Tube (translated from English under editorial supervision of V.T. Ovcharov) Izd. Sovetskoye radio, 1952, 39.
13. S.B. Cohn, IRE Trans., 1954, MTT-2, 52.

Submitted to the editors 27 June 1960

INFLUENCE OF TWO TYPES OF LONGITUDINAL PERIODIC MODULATION ON THE PROPERTIES OF A SINGLE-BEAM ELECTRON STREAM

G.F. Filimonov

The report examines the laws of conservation and the mode of weak oscillations in an electron stream for which there is known either the initial (periodic in space) distribution of electrons and their velocities along the beam or the boundary (periodic in time) values of density and velocity of electrons for all moments in time. The limits of applicability of the Euler equations are determined for the discussed processes. The results are compared with those of other authors.

INTRODUCTION

As the result of a number of investigations devoted to the study of oscillations in uniform single-velocity electron streams it may be regarded as established that for definite relationships between the parameters of the electron beam there are excited within it extremely intense electron oscillations in the direction of their initial movement. It is also known that in the process of buildup in these oscillations the logarithmic decrement of the buildup in their amplitudes drops to zero, after which the beam exhibits stable modulation.

Thus, modulation of the electron beam affects its dispersion properties, altering the regions of instability existing in uniform beams and eliminating them at a certain percentage modulation. On the other hand, modulation of an electron stream may give rise to new regions of instability which were absent in an initially uniform stream. Hence, it is of interest to study the influence of periodic modulation on the dispersion properties of an electron stream. Interest in such phenomena has recently increased, particularly due to the appearance of a number of devices directly using periodically modulated beams.

It should be mentioned that published works devoted to examination of problems of stability of periodic beams (for example, references [1-4]) contain serious procedural shortcomings, which necessitates a rigorous check of the applicability of the results obtained therein. On the other hand, simple analysis shows that different types of periodic modulation leads to substantial differences in the physical properties of the electron stream, a fact which has received little attention in the literature.

The subject of the influence of external modulation on the properties of an electron stream is extermely complex. Two aspects of this subject lend themselves to relatively simple analysis: the laws of conservation and investigation of stability by means of linearized equations. The present report is devoted to an examination of these questions for the

case of a unidimensional single-beam electron stream with strictly periodic (in time or in space modulation.

1. BASIC RELATIONSHIPS. LAWS OF CONSERVATION

Below we examine systems of charges of one sign. If such systems are not subjected to external influences, the charges within them will, generally speaking, repel one another and the various expressions for the field of the interacting charges will describe the effects of repulsion. The only exception to this rule may be those natural waves, the phase velocities of which are close to the electron velocities and which, consequently, may have prolonged accelerating or retarding effect on the charges. Such waves of "resonant interaction" with the electrons are conveniently distinguished from the total spectrum of waves in the system, since their action on the streams of charges may cause bunching of charges into bundles, which is the opposite effect of repulsion. The corresponding equation of motion in the case under discussion may be presented in the form

$$\frac{d^2z}{dt^2} = \frac{e}{m}\{E_0(z,t) + E_{\parallel}(z,t) + E_{\perp}(z,t)\}, \tag{1}$$

where z and t are the present coordinate and time of the electrons; e and m are the electron charge and mass, respectively $E_{\parallel}$, $E_{\perp}$, E_0 are the fields of the repulsive forces, of the resonance wave and a certain external field. respectively. The second basic equation for this problem is an equation of continuity giving the relationship between the linear current and linear density of the electrons.

$$\frac{\partial \rho(z,t)}{\partial t} + \frac{\partial j(z,t)}{\partial t} = 0, \tag{2}$$

The third equation is given for $E_{\parallel}$, $E_{\perp}$, expressing $E_{\parallel}$, $E_{\perp}$ in terms of ρ or j. In the case of a process which is strictly periodic in time $E_{\perp}$ and $E_{\parallel}$ are defined by the equalities (references [5, 6]

$$\begin{gathered} E_{\perp} = \operatorname{Re} E(z)\, e^{-i\omega t}, \\ \left(\frac{d^2}{dz^2} + h^2\right) E(z) = -iR_0 \int_0^T \frac{dt}{T}\, j(z,t)\, e^{i\omega t}, \\ E_{\parallel} = \frac{4\pi}{S}\int_0^T dt'\, j(z,t')\, D(\omega t' - \omega t), \\ D(x) = D(x \pm 2\pi) = \frac{\operatorname{sh} p(\pi - x)}{2 \operatorname{sh} p\pi}, \quad 0 < x < 2\pi, \end{gathered} \tag{3a}$$

where T is the time period; $\omega = 2\pi/T$; h is resonance wave number corresponding to frequency μ; p^{-1} is the effective radius of action of the repulsive forces, its value being close to $p^{-1} \simeq \pi \frac{b}{\lambda}\frac{c}{v_0}$ (c is the velocity of light, $\lambda = cT$, b is the beam radius, v_0 is the mean electron velocity during a period); S is the effective cross section of the electron beam* With $t = t(z, t_0)[t(0, t_0) = t_0]$ as the solution of (1), we may make the following change under the integral signs in (3a).

$$j(z,t)\, dt(z,t_0) \to j(0,t_0)\, dt_0, \tag{4a}$$

whereupon (1) and (3a) form a closed system of equations in the Lagrange variables z, t_0. The current $j(0, t_0)$ entering the device is considered known. Formula (4a) is equivalent to the equation of continuity (2).

---

*More rigorous formulas for calculation of p and S are given in reference [6]; formulas for calculation of R_0 are givwn in reference [5].

In the case of a process which is strictly periodic in space, for $E_\perp$ $E_\parallel$ we may derive formulas*

$$E_\perp = \operatorname{Re} E(t) e^{i\gamma z}, \quad \left(\frac{d^2}{dt^2} + \omega^2\right) E(t) = iR_0 \int_0^L \frac{dz}{L} \rho(z,t) e^{-i\gamma z}, \tag{3b}$$

$$E_\parallel = \frac{4\pi}{S} \int_0^L dz' \rho(z', t) D(\gamma z - \gamma z'),$$

similar to (3a). Here L is the spatial period: $\gamma = 2\pi/L$: ω is the natural frequency of the resonance wave, corresponding to the wave number γ; $p^{-1} \simeq \pi b/L$. With $z = z(t, z_0)[z(0, z_0) = z_0]$ as the solution of (1), in the righthand sides of (3b) we may make the following change [equivalent to the solution of (2)].

$$\rho(z, t)\, dz(t, z_0) \to \rho(z_0, 0)\, dz_0, \tag{4b}$$

whereupon the system of (1) and (3b) is closed if the initial density $\rho(z, 0)$ is known.

Let us now discuss the first integrals of motion contained in (1) for processes which are strictly periodic in time and for processes which are strictly periodic in space. Multiplying (1) first by j(z,t) and then by $\rho(z,t)$, assuming $\mathbf{E}_0 = 0$ and integrating (1) for the' period T for processes which are periodic in time, we obtain

$$\int_0^T \frac{dt}{T} \frac{j(z,t)}{e} \frac{m}{2} \left(\frac{\partial t}{\partial z}\right)^{-2} + \frac{d\varphi(z)}{dz} \frac{A^2(z)}{R_0} = \text{const},$$

$$\int_0^T \frac{dt}{T} \frac{j(z,t)}{e} m \left(\frac{dt}{\partial z}\right)^{-1} - \int_0^T\int dt\, dt' \frac{j(z,t)\, j(z,t')}{S} D_1(\omega t' - \omega t) -$$

$$- \left[\left(\frac{d^2}{dz^2} + h^2\right) \frac{A^2(z)}{2\omega R_0} - 3 \frac{|E'(z)|^2}{2\omega R_0}\right] = \text{const}, \tag{5a}$$

$$E(z) = A(z) e^{i\varphi(z)},$$

$$D_1(x) = D_1(x \pm 2\pi) = \frac{\operatorname{ch} p(\pi - x) - 1}{2 \operatorname{sh} p\pi p}, \qquad 0 < x < 2\pi.$$

The equalities in (5a) have more than once been derived in various approximations in the literature and their physical meaning is known. In the absence of a resonance wave [A(z) = E(z) = 0] they describe the movement of electrons in the self-consistent field of repulsive forces. The first and second equalities of (5a) show in this case that the mean flux of kinetic energy of electrons per period remains constant and that the energy of the field of repulsive forces rises simultaneously with the flux of the alternating component of the electron impulse. This indicates that the swing in the beam of oscillations of finite intensity is the result of natural separation of the beam electrons into two groups, one of which is retarded by the total repulsive field and the other is accelerated. A characteristic feature of such separation of the beam is the conservation of the mean electron current per period [see (2)] and of the flux of their kinetic energy [see the first equality in (5a)].

$$\overline{j(z,t)}^T = \int_0^T \frac{dt}{T} j(z,t) = \text{const},$$

$$\overline{j(z,t)\left(\frac{\partial t}{\partial z}\right)^{-2}}^T = \int_0^T \frac{dt}{T} j(z,t) \left(\frac{\partial t}{\partial z}\right)^{-2} = \text{const}. \tag{6a}$$

*In the derivation of (3b) for E || the following assumptions were made: unidimensionality of the process, strict periodicity of all quantities in z, omission of terms $\sim c^{-2}$. For precise calculation of p in (3b) all formulas and numerical results presented in reference [6] are suitable if α_k^1 therein is replaced by γ.

The appearance of the resonance wave [see first equality in (5a)] leads to systematic retardation (or acceleration) of all electrons of the beam, which is accompanied by systematic decrease (or increase) in flux of kinetic energy of the electrons and an equal increase (or decrease) in flux of energy of the resonance wave.

For processes which are periodic in space, instead of (5a), we obtain the equalities ($E_0 = 0$)

$$\int_0^L \frac{dz}{L} \frac{\rho(z,t)}{e} m \frac{\partial z}{\partial t} + \frac{d\psi}{dt} \frac{A^2(t)}{R_0} = \text{const}, \quad E(t) = A(t) e^{-i\psi(t)},$$

$$\int_0^L \frac{dz}{L} \frac{\rho(z,t)}{e} \frac{m}{2} \left(\frac{\partial z}{\partial t}\right)^2 + \int_0^L \int_0^L dz\, dz' \frac{\rho(z,t)\rho(z',t)}{S} D_1(\gamma z - \gamma z') - \tag{5b}$$

$$- \left[\left(\frac{d^2}{dt^2} + \omega^2\right) \frac{A^2(t)}{2\gamma R_0} - 3 \frac{|E'(t)|^2}{2\gamma R_0}\right] = \text{const},$$

which are quite similar to those of (5a) and have the same physical meaning. With A(t)=0 it follows from these equalities that the mean electron current is conserved and that the energy of the repulsive field may increase only due to a decrease in the mean electron energy per period. Here, as in the case of processes which are periodic in time, swinging of oscillations of finite amplitude is the consequence of division of the initial beam into two groups of electrons moving in opposite phase. However, the existence of a direct transition of the kinetic energy of electrons into the energy of the repulsive field appears to attest to the excitation within spatially periodic systems of space-charge waves. The mean current [see the first equality in (5b)] and mean charge density [see (2)] remain constant in the presence of spatially periodic oscillations of electrons in the repulsive field:

$$\overline{j(z,t)}^L = \int_0^L \frac{dz}{L} j(z,t) = \text{const}, \qquad \overline{\rho(z,t)}^L = \int_0^L \frac{dz}{L} \rho(z,t) = \text{const}. \tag{6b}$$

The resonance wave, as in the case of processes which are periodic in time, causes a systematic change in electron velocity and mean current. With $E_0 \neq 0$ and $E_\perp = 0$ the quantities $\overline{j(z,t)}^L$ and $\overline{j(z,t)\left(\frac{dt}{dz}\right)}^T$ are not preserved.

2. LINEARIZATION OF EQUATIONS

Derivation of analytical solutions of (1) in our case of electron movement within the field of natural repulsive forces ($E_0 = E_\perp = 0$) does not appear possible at this time. Hence, below we examine solutions of the linearized equation (1) in the limiting cases of long-wave and short-wave disturbances. On the assumption that the modulation parameters are small, let us linearize (1). In operator form this is equivalent to the conditions

$$\left|\frac{\partial}{\partial \omega t}\right| \ll 1, \quad \left|\frac{\partial}{\partial \gamma z}\right| \ll 1$$

Then, using (2) and the expression for current $j = \rho v$, instead of (1) we obtain

$$\left(\frac{\partial}{\partial t} + v_0 \frac{\partial}{\partial z}\right) j(z,t) - v_0 \left(\frac{\partial}{\partial t} + v_0 \frac{\partial}{\partial z}\right) \rho(z,t) = \rho_0 \frac{e}{m} E_\parallel. \tag{7}$$

Here $v_0 = v_0(z,t)$, $\rho_0 = \rho_0(z,t)$ are certain zero-th approximations of $v(z,t)$, $\rho(z,t)$ around which (1) is linearized. In the general case, by v_0 and ρ_0 we must understand a certain traveling wave or superposition of traveling waves. It is easily seen that in the case of a traveling wave, when the unique argument of ρ_0 v_0 is the combination $\gamma z - \omega z$ by linear transformation of variables z,t we may reduce ρ_0, v_0 to a function of only one variable t' to z'. Then equations (7) and (3) will yield (in the moving coordinate system of coordinates) solution of the problem with given boundary values [$\rho_0 = \rho_0(t')$, $v_0 = v_0(t')$] or with given initial values ($\rho_0 = \rho_0(z')$, $v_0 = v_0(z')$, respectively. It is not difficult to see that by virtue

of the approximate character of equation (7) these solutions will be valid only within the limits of a narrow interval of the oz axis or the ot axis adjacent (in the moving coordinate system) to the point of application of the given boundary or initial values, respectively.*

Using the proximity of the required solution of (7) to an absolute constant (weak modulation), we may consider ρ_0, v_0 to be the function of a single variable (t or z) as well as directly in a laboratory system of coordinates. From the mathematical point of view this is equivalent to a certain transposition of terms of the first (relative to the modulation parameter) order of smallness between the solution of (7) and its zero-th approximation, while from the physical point of view such an approach permits investigation of the influence of a given modulation in a certain plane z = 0 or a given time t = 0 on beam stability. Since the second approach to solution of (7) leads to simpler expressions and proves applicable both in the case of solution of (7) in the vicinity of a single traveling wave and in the case of solution close to any superposition of traveling waves, below we shall consider $\rho_0 v_0$ as functions only of t or z. However, it must be kept in mind that the resulting solutions are suitable for different space-time regions.

Even with these simplifying assumptions Equation (7) does not have a simple solution in view of the integral nature of the $E_\parallel$ field [see (3)]. Further simplification of (7) may be achieved in the limiting cases of strong ($p \to \infty$) and weak ($p \to 0$) shielding of the space-charge field generated by a given cross section of the electron beam. The case $p \to \infty$ corresponds to thin beams or long-wave perturbations $\left(b \ll \lambda \frac{v_0}{c},\ b \ll L\right)$. With series expansion of $E_\parallel$ in powers of p^{-1} we obtain by approximation

$$E_\parallel = \frac{4\pi}{Sp^2\omega^2} \frac{\partial j(z,\ t)}{\partial t} \qquad j(z,\ t) \equiv j(z,\ t+T), \tag{8a}$$

$$E_\parallel = -\frac{4\pi}{Sp^2\gamma^2} \frac{\partial \rho(z,\ t)}{\partial z}, \qquad \rho(z,\ t) \equiv \rho(z+L,\ t). \tag{8b}$$

The case $p \to 0$ corresponds to wide beams or short-wave disturbances $\left(b \gg \lambda \frac{v_0}{c},\ b \gg L\right)$. Expansion in powers of p yields by approximation

$$\frac{\partial E_\parallel}{\partial t} = -\frac{4\pi}{S}\left[j(z,\ t) - \overline{j(z,\ t)}\right],$$
$$\frac{\partial E_\parallel}{\partial z} = \frac{4\pi}{S}\left[\rho(z,\ t) - \overline{\rho(z,\ t)}\right]. \tag{9}$$

Averaging in (9) must be performed for the period T(L) if the process is periodic in time (in space).

Above we have seen that conservation of the mean values of current per period and density per period is possible only with strict periodicity of processes in time or in space. Hence the second terms of equalities (9), which do not lead to complications in the case of strictly periodic processes, may give rise to a self-maintaining branch of oscillations with intrinsic regions of instability in the case of nonperiodic (or almost nonperiodic) processes.

3. LONG-WAVE DISTURBANCES

In this case, inserting, for example, (8b) into (7) and eliminating therefrom $\rho(z,t)$ by means of (2), we obtain for j(z,t) a second-order partial differential equation of the

*This circumstance may be explained by the following rough evaluation. Exact solution of Equations (1), (2), [(3a) or (3b)] is given by the functions $v=v(z,t), \rho=\rho(z,t)$ defined over the entire plane z,t. In the process of linearization we distinguish in plane z,t a certain line $z=z_0(t)$ at which the function $v[z_0(t), t] = v_0[z_0(t), t]\rho[z_0(t)]=\rho_0[z_0(t), t]$ is considered known and Equations (1)-(3) undergo series expansion in powers of the quantities $\Delta v = v - v_0$, $\Delta\rho = \rho - \rho_0$, which we consider to be small. The resulting linearized equations yield solutions which are approximately accurate only in the vicinity of the fixed line $z = z_0(t)$. In practice such a line is usually represented by the time axis (the problem with boundary values v_0, ρ_0, given in $z=z_0=\text{const}$) or the coordinate axis (the problem with initial values of values of $v_0\rho_0$ given at the moment $t=t_0=\text{const}$).

the hyperbolic type. In the space of characteristics it is described as follows:

$$\frac{\partial^2}{\partial X \partial Y} \rho_0^{-1/4}(z)\, j(z,\ t) - j(z,\ t) \frac{\partial^2}{\partial X \partial Y} \rho_0^{-1/4}(z) = 0, \tag{10}$$

where variables X, Y are associated with z, t by the equalities

$$X = t - \int_0^z \frac{dz}{v_0(z) + \dfrac{\omega_0(z)}{\gamma p}},$$

$$Y = \int_0^z \frac{dz}{v_0(z) - \dfrac{\omega_0(z)}{\gamma p}} - t, \quad \omega_0^2(z) = \frac{4\pi e \rho_0(z)}{mS}. \tag{11}$$

Equation (10) is solved by the Riemann method (e.g., see reference [7]). In the first approximation relative to the modulation parameter its solution is

$$j(z,\ t) = \frac{j_0(z_2) + j_0(z_1)}{2} + \frac{p^2}{1+p^2} \frac{\omega_{00}}{\gamma p} \frac{\rho_0(z_2) - \rho_0(z_1)}{2} + j_{00} \frac{\gamma p}{\omega_{00}} \frac{v_0(z_2) - v_0(z_1)}{2}, \tag{12}$$

where

$$\omega_{00}^2 = \overline{\omega_0^2(z)}^L;\ j_{00} = \overline{j_0(z)}^L; \int_{z_1}^z \frac{dz}{v_0(z) - \dfrac{\omega_0(z)}{\gamma p}} = \int_{z_2}^z \frac{dz}{v_0(z) + \dfrac{\omega_0(z)}{\gamma p}} = t;$$

$z_{1,2}$ are easily calculated in the first approximation relative to the modulation parameter. In this approximation the Riemann function is

$$U(X,\ Y;\ \xi,\ \eta) = \rho_0^{-1/4}(z(X+Y))\, \rho_0^{1/4}(z(\xi+\eta))\, \rho_0^{1/4}(z(\xi+Y))\, \rho_0^{1/4}(z(X+\eta)).$$

For the time-periodic process, solution of (7) is similarly found and proves to be

$$\rho(z,\ t) = \frac{\rho_0(t_1) + \rho_0(t_2)}{2} + \frac{p^2}{1+p^2} \frac{\omega_{00}}{\omega v_{00} p} \frac{j_0(t_1) - j_0(t_2)}{2} - \\ - \frac{\omega \rho_{00}}{\omega_{00} v_{00}} p \frac{v_0(t_1) - v_0(t_2)}{2}, \tag{13}$$

where

$$\rho_{00} = \overline{\rho_0(t)}^T;\ v_{00} = \overline{v_0(t)}^T;\ \omega_{00}^2 = \overline{\omega_0^2(t)}^T; \int_{t_1}^t \frac{v_0(t)\, dt}{1 + \dfrac{\omega_0(t)}{p\omega}} = \int_{t_2}^t \frac{v_0(t)\, dt}{1 - \dfrac{\omega_0(t)}{p\omega}} = z.$$

If we now express (12) and (13) as Fourier integrals in terms of t and z, respectively, we easily see that all the natural frequencies of (12) and the natural wave numbers of (13) are real and that the solutions of (12) and (13) belong to the class of steady-state solutions. However, when $p \to \infty$ the last term of these solutions increases linearly with p. Presenting v_0, ρ_0 in the form

$$v_0 = v_{00}(1 + \alpha \cos \Omega),\ \rho_0 = \rho_{00}(1 + \beta \cos(\Omega + \varphi)), \tag{14}$$

$\Omega = \gamma z$ for spatially periodic processes and $\Omega = \omega t$ for time-periodic process, we find that each of the terms of (12) and (13) is of the nature of an amplitude-modulated signal with a beat period

$$\Delta t \simeq \frac{\pi p}{\omega_{00}},\ \Delta z \simeq \frac{\pi p}{\omega_{00}} v_{00}$$

respectively. Hence it follows that the amplitude of long-wave modulation in the beam is determined by the velocity-modulation percentage and by the beat period and that it is always finite.

4. SHORT-WAVE DISTURBANCES

This case has often been discussed in the literature (e.g., see references [1-4]) and reduces to solution of the system of (7) and (9), which, by eliminating j(z,t) or (z,t) from (7) and omitting the squared (modulation parameter) terms, may be represented in the form

of one of the following second-order equations:

$$\left(\frac{\partial}{\partial z} + \frac{1}{v_0}\frac{\partial}{\partial t}\right)^2 j(z,t) = -\frac{4\pi e\rho_0}{mSv_0^2}[j(z,t) - \overline{j(z,t)}], \tag{15a}$$

$$\left(\frac{\partial}{\partial t} + v_0\frac{\partial}{\partial z}\right)^2 \rho(z,t) = -\frac{4\pi e\rho_0}{mS}[\rho(z,t) - \overline{\rho(z,t)}]. \tag{15b}$$

References [1-4] investigate the conditions of appearance of instability due to a periodic external force acting at point z = 0. For this purpose these studies use Equation (15a) with the z-dependent periodic coefficients. Its solution is sought in the form of a time-periodic function.

Such a statement of the problem is erroneous. Equations (15) with z-dependent coefficients are obtained by linearization of (1) around the initial values of ρ and v and yield a linear approximation which is valid only for a sufficiently small interval of time. The strictly time-periodic solution of (1) is obtained with $t \to \infty$, when we may completely disregard the influence of initial distribution. Hence, the natural wave numbers in references [1-4] are known to be derived in the presence of spreading of local discontinuities (caused by an external force) in an electron beam with initial periodic distribution of electrons in length, but these numbers have no direct relation to the decrements in the buildup of periodic disturbances close to the boundary values of ρ, v.

A second shortcoming of the work in references [1-4] is the fact that solution of (15a) is sought by expansion of the Fourier integral of j(z,t) in terms of variable t. In addition, for $j(z,\omega)$ the Hill equation is obtained with ω-dependent coefficients. Since j(z,t) is represented in this case by a contour integral $j(z,\omega)e^{i\omega t}$ in the complex plane, but the literature contains no information concerning the analytical aspects and specific features of solutions of the Hill equation with arbitrary complex coefficients, calculation of this integral by any reasonable method and derivation of an explicit expression for j(z,t) are not possible.

In order to avoid these difficulties the equations in (15) are best solved directly by describing them in the characteristic space. Then, using (15a) and (15b) for description of time-periodic and space-periodic processes, respectively, we obtain instead of (15) ordinary second-order differential equations with periodic coefficients. Schematic representation of these equations is given in the form

$$\left[\frac{d^2}{d\tau^2} + \Psi(\tau)\right] y(\tau) = 0, \quad \Psi(\tau) \equiv \sum_{m=-\infty}^{\infty} \Theta_{|m|}\cos(2|m|\tau - \varepsilon_{|m|}), \tag{16}$$

where

$$y = j - \overline{j}^T; \quad 2\tau = \omega\frac{t_1 + z_1}{v_{00}}\left(d\tau \equiv \frac{\omega}{2v_{00}}\,dz_1\right); \quad z = z_1;$$
$$\int_0^t v_0(t)\,dt = t_1 + z_1; \quad \Psi(\tau) = \frac{16\pi e\rho_0\left(t\left(\frac{2\tau v_{00}}{\omega}\right)\right)v_{00}^2}{mSv_0^2\left(t\left(\frac{2\tau v_{00}}{\omega}\right)\right)\omega^2} \tag{17a}$$

in the case of Equation (15a) and

$$y = \rho - \overline{\rho}^L; \quad 2\tau = \gamma v_{00}(t_1 + z_1)\left(d\tau \equiv \frac{\gamma v_{00}}{2}\,dt_1\right); \quad t = t_1;$$
$$\int_0^z \frac{dz}{v_0(z)} = t_1 + z_1; \quad \Psi(\tau) = \frac{16\pi e\rho_0\left(z\left(\frac{2\tau}{\gamma v_{00}}\right)\right)}{mS(\gamma v_{00})^2} \tag{17b}$$

in the case of equation (15b).

According to general theory, with θ_0 in the vicinity of the squares of integers ($\theta_0 \simeq N^2$) the solutions of Equation (16) are unstable and contain exponential coefficients $\exp[\pm\mu_N\tau]$. Width of the region of instability of N-th resonance (i.e., the range of values of $|\theta_0 - N^2|$ within which μ_N is real) and the values of μ_N prove to be quantities $\sim\theta_N$ which, in turn, are proportional to α^N, β^N [see (14), (16), (17)]. Calculation of these for N > 1 is meaningless, for the entire calculation is performed with accuracy only up to effects of the first order of smallness. For N = 1

$$\mu_1 = \frac{\Theta_1}{2}\sqrt{1 - \left(\frac{\Theta_0 - 1}{\Theta_1}\right)^2} + O(\Theta_1^2). \tag{18}$$

It is seen from (18) that μ_1 has its maximum value with $\theta_0 = 1$. By means of formulas (17) we may write the explicit expressions of θ_0, θ_1 for time-periodic processes (see(14)):

$$\Theta_0 = \left(\frac{2\omega_{00}}{\omega}\right)^2,$$
$$\Theta_1 = \frac{\Theta_0}{2}\sqrt{\beta^2 + 4\alpha^2 - 4\alpha\beta\cos\varphi}, \tag{19a}$$

and space-periodic processes:

$$\Theta_0 = \left(\frac{2\omega_{00}}{\gamma v_{00}}\right)^2, \quad \Theta_1 = \frac{\Theta_0}{2}\beta. \tag{19b}$$

The first singularity of formulas (19) lies in the fact that in the case of time-periodic processes the occurrence of instability is determined by the external-force parameter ω, while in the case of space-periodic processes it is determined by the initial-distribution parameter γ. In this connection it is of interest to point out that in real beams, along with the fundamental oscillations with period T and L, there also exist all harmonics $T_n = \frac{T}{n}(\omega_n = n\omega)$, $L_n = \frac{L}{n}(\gamma_n = n\gamma)$. Since the decrement in the buildup of μ_N is maximum for N = 1, for the condition $(2\omega_{00}/\omega_n)^2 \simeq 1$ or $(2\omega_{00}/\gamma_n v_{00})^2 \simeq 1$ it is evident that the n-th harmonic will increase and not the fundamental.

Another distinctive feature of formulas (19) lies in the fact that the decrement in buildup of time-periodic processes depends both on the density modulation percentage and on the velocity modulation percentage, whereas the decrement in buildup of space-periodic processes is determined only by the density modulation percentage.

It is of considerable interest to calculate μ_1 in the case where the preliminary modulation is close to one of the natural waves of the electron stream and the quantities v_{00}, ρ_{00}, ω, α, β, γ, φ are all associated by the initial equations (1)-(3). Assuming in (14) that $\Omega = \gamma z - \omega t$ and inserting (14) into (1)-(3), with an accuracy of the first order of smallness we find

$$\varphi = \pi m, \quad \beta = \alpha\frac{\gamma v_{00}}{\omega_{00}}\sqrt{1 + p^2}, \quad \omega = \gamma v_{00} + \omega_{00}\frac{(-1)^m}{\sqrt{1 + p^2}}. \tag{20}$$

It it is considered that in the case under discussion $p \to 0$ and $\theta_0 \simeq 1$, then it develops that for the time-periodic process the maximum decrement in buildup ($\theta_0 = 1$) of a fast wave (m = 1) is $\mu_1 = 5\alpha/4$ and is five times greater than the decrement in buildup of a slow wave if their velocity modulation parameters are identical. If their density modulation parameters are identical, then the maximum decrement in buildup of the fast wave is $\mu_1 = 5\beta/12$ and is only 1.66 time greater than the decrement in buildup of the slow wave.

In the case of space-periodic processes the maximum decrement in buildup of the fast wave is equal to the decrement in buildup of the slow wave if their density modulation parameters are identical, and is three times greater than the baximum decrement in buildup of the slow wave ($\mu_1 = \alpha/4$) in the case of identical velocity modulation.

Comparison of formulas (18), (19a), (19b) with similar results in references [1-4] is best achieved with the example of reference [2]. Formula (21) in reference [2] yields for the decrement of spatial buildup our expression (18) with the following values of coefficients θ_0 and θ_1 (given in reference [2] as c_0 and c_1)

$$\Theta_0 = \left(\frac{2\omega_{00}}{\gamma v_{00}}\right)^2, \quad \Theta_1 = \alpha(\beta - \Theta_0). \tag{19c}$$

Reference [2] examines the case of $j_0(z) = \text{const}$ (i.e., $\alpha = \beta$, $\varphi = \pi$). From formulas (19) it is seen that the regions of swing derived in reference [2] for time-periodic oscillations generated by a time-periodic external force applied at moment t = 0 (when the beam has a spatially periodic structure) are determined by parameter γ and coincide with the regions of instability of the spatially periodic stream which exist also in the absence of an external

force and have nothing in common with the regions of instability of the time-periodic beam which are determined by the external-force parameter ω. The numerical values of μ_1 obtained from formulas (18) and (19c) also do not coincide with the values of μ_1 defined in (18), (19a), (19b).

Hence, it follows that the type of instability investigated in references [1-4] is not attained with strict spatial or temporal periodicity of oscillations of electrons in the beam, as might have been assumed at the very outset. The general solution of Equation (16) has the form

$$y(\tau) = y(\tau_0)\frac{y_1(\tau)\,y_2'(\tau_0) - y_2(\tau)\,y_1'(\tau_0)}{y_1(\tau_0)\,y_2'(\tau_0) - y_2(\tau_0)\,y_1'(\tau_0)} + y'(\tau_0)\frac{y_1(\tau_0)\,y_2(\tau) - y_1(\tau)\,y_2(\tau_0)}{y_1(\tau_0)\,y_2'(\tau_0) - y_2(\tau_0)\,y_1'(\tau_0)}, \tag{21}$$

where $\tau_0 = \omega t/2v_{00}$ in the case of (17a) and $\tau_0 = \gamma v_{00} z/2$ in the case of (17b). The partial solutions of $y_{1,2}(\tau)$ are conveniently presented in the form of series in powers of small parameters α, β. Since $y(\tau)$, $y'(\tau_0)$ are quantities of the first order of approximation[see (17a) (17b)], for $y_k(\tau)$ it is necessary to take the zero-th approximation in powers of α, β. It may be shown that with $|\sqrt{\theta_0} - m| \gg \theta_1$, $\pm m = 0, 1, 2, \ldots$ the zero-th approximation of steady-state partial solutions of (16) within the region of absolute convergence defined by the inequality*

$$\left|\frac{\theta_1}{4(N-\sqrt{\theta_0})}\left[\frac{2N+2+N(N-\sqrt{\theta_0})}{(N+2)(\sqrt{\theta_0}+1)} + \frac{2N-2-N(N-\sqrt{\theta_0})}{(N-2)(\sqrt{\theta_0}-1)}\times\right.\right.$$

$$\left.\left.\times\frac{1+\frac{N-2,5}{|N-2,5|}}{2}\right] + O(\theta_1^2)\right| < 1, \tag{22}$$

are

$$y_1 = \sin\sqrt{\theta_0}\,\tau,\quad y_2 = \cos\sqrt{\theta_0}\,\tau. \tag{23}$$

The corresponding solutions of (16) have the form

$$j(z,t) - j_{00} = \left[j_0\left(t - \frac{z}{v_{00}}\right) - j_{00}\right]\cos\frac{\omega_{00}z}{v_{00}} + \rho_{00}\frac{\partial v_0\left(t - \frac{z}{v_{00}}\right)}{\partial t}\frac{\sin\frac{\omega_{00}z}{v_{00}}}{\omega_{00}}, \tag{24a}$$

$$\rho(z,t) - \rho_{00} = [\rho_0(z - v_{00}t) - \rho_{00}]\cos\omega_{00}t - \rho_{00}\frac{\partial v_0(z - v_{00}t)}{\partial z}\frac{\sin\omega_{00}t}{\omega_{00}} \tag{24b}$$

for time-periodic processes (24a) and space-periodic processes (24b).

Keeping in mind that in the first approximation of the transient solution it is necessary to consider $e^{\pm\mu 1\tau} = 1$ it may be shown that it coincides in accuracy with (24) if we consider in (24) the equality $\theta_0 = 1$.

Formula (24a) is derived on the same assumptions as (22) (reference [8]) and in principle has the same area of applicability. However, in inserting (24a) into (1), obtaining the approximate solution and comparing the result

$$t = t_0 + \frac{z}{v_{00}} + \frac{\alpha\sin\omega t_0 + \beta\sin(\omega t_0 + \varphi)}{\omega}\left(1 - \cos\frac{\omega_{00}z}{v_{00}}\right) - \frac{\alpha}{\omega_{00}}\cos\omega t_0\sin\frac{\omega_{00}z}{v_{00}} \tag{25}$$

with (22) [8], we find that (25a) does not coincide with (22) [8] within the first order of smallness. This deviation is apparently due to the fact that in deriving basic equation (18) in reference [8] there is made a change in $v^2 \to v_{00}^2$ which has unverifiable results.

5. VALIDITY OF EULER EQUATIONS

Basic linearized equation (7) is written in the Euler form and is applicable to those space-time regions where the velocity field is single-valued, continuous and differentiable. Since these conditions are disturbed as the electrons overtake one another, the Euler equations must be considered valid only in absence of such behavior. Formula (25) permits evaluation of the region of variation in parameters within the limits of which the overtaking may

*In order of magnitude, condition (22) is equivalent to $\alpha, \beta < \frac{\text{const}}{\sqrt{\theta_0}}$.

be disregarded.

The effect of overtaking consists in the fact [e.g., see (25)] that to one t, z there correspond several values of t_0. Hence the values of z at which overtaking occurs may be found from the equation

$$\frac{\partial t(z, t_0)}{\partial t_0} = 0 \tag{26}$$

in the case of time-periodic processes. Choosing t_0 so that the root of (26) is minimum, we obtain the point of the first overtake. With the requirement that the characteristic range of solution of (24a) which are equal to $2\pi v_{00}/|\omega \pm \omega_{00}$, do not exceed z_{ov}, we find the following conditions of applicability of the Euler equations:

$$\omega_{00} > \alpha\omega;\ \omega_{00} < \alpha\omega,$$

$$z_{ov} = \frac{v_{00}}{\omega_{00}} \arcsin \frac{\omega_{00}}{\alpha\omega} > \frac{2\pi v_{00}}{|\omega \pm \omega_{00}|} \tag{27a}$$

in the case of time-periodic processes. With $\omega_{00} > \alpha\omega$ Equation (26) has no solution (that is, overtaking is absent).

By repeating the reasoning given above we may show that space-periodic processes are described by the Euler equations upon fulfilment of the conditions

$$\omega_{00} > \alpha\gamma v_{00};\quad \omega_{00} < \alpha\omega,$$

$$t_{ov} = \frac{1}{\omega_{00}} \arcsin \frac{\omega_{00}}{\alpha\gamma v_{00}} > \frac{2\pi}{|\omega \pm \omega_{00}|}. \tag{27b}$$

It is seen from (27) that oscillations without overtaking occur only in sufficiently dense electron beams. With small density the overtaking may be disregarded only with sufficiently low velocity modulation [in order of magnitude the second inequalities in (27a) and (27b) are equivalent to $\alpha > 1/2\ \pi$]. In the literature on klystron theory ($\beta = 0$) expressions have already been derived for z_{ov}. It is of interest to note that density modulation of the beam causes a shift of the first point of overtaking by only an infinitesimally small magnitude of the first order.

In the case of transient oscillations the criterion of applicability of the Euler equations does not have the form of (27). However, its valid derivation requires the solution of nonlinear equations taking into account effects of the second order of smallness.

REFERENCES

1. O.E.H. Rydbeck, B. Agdur, Onde electr. 1954, 34, 499.
2. P.V. Bliokh, Ya.V. Faynberg, ZhTF, 1956, 26, 530.
3. Yu.F. Fillippov, Radiotekhnika i elektronika, 1959, 4, 2, 228.
4. V.A. Solntsev, A.S. Tager, Radiotekhnika i elektronika, 1959, 4, 10, 1652.
5. L.A. Vaynshteyn, ZhTF, 1953, 23, 654.
6. G.F. Filimonov, Radiotekhnika i elektronika, 1959, 4, 3, 489.
7. F. Trikomi, Lectures on equations in partial derivatives, IL, 1957,
8. V.Ya. Savel'yev, ZhTF, 1940, 10, 1865.

Submitted to the editors 29 June 1960

ON A THEORY OF DIODE MICROWAVE OSCILLATORS

V.A. Malyshev

Within the framework of kinematic approximation the report presents a theory of diode microwave oscillators and an analysis of their operation. A comparison of diode oscillators and reflex klystrons is made on the basis of their output parameters.

1. INTRODUCTION. ELECTRON ADMITTANCE

Information contained in the literature on the theory of diode oscillators (references [1,2,3,4]) is limited to a study of the physics of processes in the diode and devotes little attention to the specific features of operation of such devices and design calculations.

Discussion of these problems is now occasioned by the definite interest in diode-type oscillators, particularly in the oscillator with retarding field [5,6,7].

The present report presents an attempt to analyze the operation of diode oscillators in small-amplitude approximation disregarding the influence of space charge on bunching processes.

The equation of motion of the electron entering the bunching space, in which there acts constant (E_0) and alternating electric field,

$$\frac{m}{\varepsilon}\ddot{x} = E_0 + E_1 \sin \omega t \tag{1}$$

with the initial conditions $t = \tau$, $x = 0$, $\dot{x} = v_0 = \sqrt{2\frac{\varepsilon}{m}U_0}$ has the solution

$$x = \left[\frac{\varepsilon}{m}\frac{E_0}{2}(t-\tau)^2 + v_0(t-\tau)\right] + \frac{\varepsilon E_1}{m\omega^2}[\omega(t-\tau)\cos\omega\tau - \sin\omega t + \sin\omega\tau]. \tag{2}$$

The first term to the right in (2) may be expressed as $(t - \tau)\bar{v}$, where the mean velocity for the case of motion to complete cessation in the retarding field is $\bar{v} = v_0/2$ and for the case of accelerated motion is

$$\bar{v} = \frac{v_0}{2}\left(1 + \sqrt{\frac{U_2}{U_0}}\right), \tag{3}$$

Where, U_2 and U_0 are the potentials of the electrodes forming the bunching space. From expression (2) we may obtain

$$\varphi_0 = \varphi - \mu A \sin(\varphi - \psi - \theta), \tag{4}$$

where

$$\varphi = \omega(t-\tau);\ \varphi_0 = \frac{\omega x}{\bar{v}};\ \psi = \omega t;\ \mu = \frac{\varepsilon E_1}{m\omega\bar{v}};$$

$$A = \sqrt{(\varphi - \sin\varphi)^2 + (1-\cos\varphi)^2};\ \operatorname{tg}\theta = \frac{\varphi - \sin\varphi}{1-\cos\varphi}.$$

In order to determine the current at the point with corrdinate x : $i = I_0\left(1 - \frac{\partial\varphi}{\partial\psi}\right)$ it is necessary to find the dependence $\varphi = f(\varphi_0, \psi)$. This dependence may be determined from (4) in the case of small-signal approximation, when $\mu \ll 1$. In addition, in the expression for A we may, without significant error, insert φ_0 instead of φ. Moreover, angle θ varies with a change in φ, so that in the expression for θ we may also replace φ with φ_0 (it may be shown that in the more interesting sectors of change $2\pi < \varphi < 3\pi$ the angle θ varies by less than 7% of

the change in φ). After such substitutions expression (4) acquires the form of the Kepler equation for φ, the solution of which will be (reference [8])

$$\varphi = \varphi_0 + \\ + \sum_{n=1}^{\infty} \frac{2}{n} J_n(n\mu A) \sin[n(\varphi_0 - \psi - \theta)] \tag{5}$$

whence for the current we obtain the relationship

$$i = I_0 \left\{1 + 2\sum_{n=1}^{\infty} J_n(n\mu A) \times \right. \\ \left. \times \sin\left[\frac{\pi}{2} + n(\psi - \varphi_0 + \theta)\right]\right\}, \tag{6}$$

from which it is easy to find the first harmonic of current.

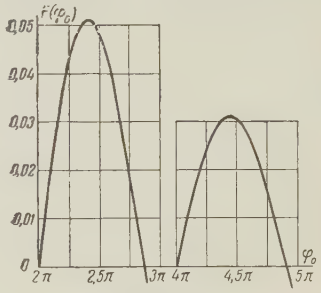


Figure 1

It must be mentioned that formula (6) may also be arrived at by series expansion of φ in expression (4) in powers of μ with subsequent determination of the current in the form of a Fourier series (references [2,3,4]).

The power of interaction of the stream with the field (P), associated with the electron admittance of the interaction sector (Y_e) by the relationship $2P = Y_e U^2$ (where U is the amplitude of voltage at the electrodes: $U = E_1 d$) may be found from the expression

$$P = \frac{1}{2}\int_0^l iE^* dx = \frac{1}{2}\frac{\bar{v}}{\omega}\int_0^{\varphi_0} iE^* d\varphi_0. \tag{7}$$

Moreover, for Y_e we obtain the formula

$$Y_e = \frac{2\bar{v}I_0}{\omega U d} e^{j\frac{\pi}{2}} \int_0^{\varphi_0} J_1(\mu A)\, e^{j(\theta - \varphi_0)}\, d\varphi_0, \tag{8}$$

wherein $A = \sqrt{p^2 + q^2}$, $\operatorname{tg}\theta = p/q$, $p = \varphi_0 - \sin\varphi_0$, $q = 1 - \cos\varphi_0$. Exact evaluation of the integral in (8) does not seem possible. Hence, by expanding the Bessel function into a series and limiting ourselves to two terms of the series, we may obtain

$$Y_e = \frac{2\bar{v}I_0}{\omega U d} e^{j\frac{\pi}{2}} \left\{\frac{L\mu}{2} e^{j\gamma} - \frac{\mu^3}{16}\int_0^{\varphi_0} (p^2 + q^2)[f(\varphi_0) + j\lambda(\varphi_0)]\, d\varphi_0\right\},$$

where

$$L = \left\{\left[\int_0^{\varphi_0} f(\varphi_0)\, d\varphi_0\right]^2 + \left[\int_0^{\varphi_0} \lambda(\varphi_0)\, d\varphi_0\right]^2\right\}^{1/2}; \quad \gamma = \operatorname{arc\,tg} \frac{\int_0^{\varphi_0} \lambda(\varphi_0)\, d\varphi_0}{\int_0^{\varphi_0} f(\varphi_0)\, d\varphi_0}; \tag{9}$$

$$f(\varphi_0) = \cos\varphi_0 - 1 + \varphi_0 \sin\varphi_0;$$
$$\lambda(\varphi_0) = \varphi_0 \cos\varphi_0 - \sin\varphi_0.$$

For evaluation of the integral in expression (9) let us use the mean value theorem and carry $p^2 + q^2$ beyond the integral sign, taking the value of φ_0 in this quantity equal to the mean $\bar{\varphi}_0$ (the relation between φ_0 and $\bar{\varphi}_0$ will be found later). Then we obtain

$$Y_e = \frac{2I_0\bar{v}L}{\omega U d A} e^{j\left(\frac{\pi}{2}+\gamma\right)}\left[\frac{X}{2} - \frac{X^3}{16}\right] \simeq \frac{2I_0\bar{v}L}{\omega U d A} J_1(\lambda)\, e^{j\left(\frac{\pi}{2}+\gamma\right)}, \tag{10}$$

where $X = \mu\overline{A}$ and $\overline{A} = \sqrt{\overline{p^2} + \overline{q^2}}$, the lines above p,q and A indicating averaging. The Bessel function appeared in expression (10) due to replacement of the two terms representing the terms of its series expansion. It may be shown that if in obtaining (9) from (8) the Bessel function had been replaced by a large number of terms of expansion, all these terms in the described process of averaging would have appeared in expression (10), resulting in a series representing the Bessel function.

After integration and substitution, the final expression for Y_e has the form

$$Y_e = 2G_0 \frac{J_1(X)}{X} e^{-j(\delta+\pi)}, \tag{11}$$

where

$$G_0 = \frac{I_0}{U_0 B^2} N(\varphi_0);$$

$$X = \frac{UM(\varphi_0)}{2U_0 B_0^2}; \qquad M(\varphi_0) = \frac{\overline{A}}{\varphi_0};$$

$$N(\varphi_0) = \frac{1}{\varphi_0}\left[\frac{\sin\frac{\varphi_0}{2}}{\frac{\varphi_0}{2}} - \cos\frac{\varphi_0}{2}\right]; \quad \delta = -\frac{3\pi}{2} + \frac{\varphi_0}{2},$$

wherein the quantities B, B_0 and φ_0 for different types of oscillators have the values; $B=B_0=1$, $\varphi_0 = \omega d/v_0$ for oscillators of the monotron type, in which the constant field is lacking in the bunching space; $B = B_0 = (\sqrt{U_0} + \sqrt{U_2})/2\sqrt{U_0}$, $\varphi_0 = 2\omega d\sqrt{U_0}/v_0(\sqrt{U_0} + \sqrt{U_2})$ for oscillators with accelerating field in the bunching space; in the case where this region begins at the cathode: $B = B_0 = 1/2$, $\varphi_0 = \omega d\sqrt{2\frac{m}{eU_2}}$ (wherein in expressions (11), instead of U_0, it is necessary to write U_2 throughout); the above consideration is valid in this case only on the condition of diode operation at saturation. Finally, for the case of the retarding field

$$2B_0^2 = B = \frac{U_0 - U_2}{4U_0}, \qquad \varphi_0 = \frac{4\omega d U_0}{v_0(U_0 - U_2)}.$$

Oscillation occurs in regions where $N(\varphi_0)/\cos\delta > 0$, and since $\cos\delta = -\sin\frac{\varphi_0}{2}$, then the condition for possibility of oscillation may be represented as

$$F(\varphi_0) = N(\varphi_0)\cos\delta = \frac{1}{2}\left[\frac{\sin\varphi_0}{\varphi_0} - \left(\frac{\sin\frac{\varphi_0}{2}}{\frac{\varphi_0}{2}}\right)^2\right] > 0. \tag{11a}$$

A plot of the relationship $N(\varphi_0)\cos\delta = F(\varphi_0)$ for the first two zones of oscillation is given in Figure 1. It follows therefrom that oscillation is possible with $2\pi < \varphi_0 < 2.86\pi$, $4\pi < \varphi_0 < 4.92\pi, \ldots$.

Within these limits of change in φ_0 the result of the above averaging must yield minimum error. This permits choice of the relationship between quantities φ_0 and $\overline{\varphi}_0$. As the result of this choice, $\overline{\varphi}_0 = \varphi_0$ for the first and second zones of oscillation. It may be shown that in this case the function $M(\varphi_0)$ is defined by the formula

$$M(\varphi_0) = \sqrt{1 + \frac{2}{\varphi_0^2}(1 - \cos\varphi_0 - \varphi_0\sin\varphi_0)}. \tag{12}$$

Due to the fact that in averaging the quantity $p^2 + q^2$ for various functions of $f(\varphi_0)$ and $\lambda(\varphi_0)$ one and the same value of $\overline{\varphi}_0$ was assumed in this quantity, the results of such averaging may show significant errors. In order to evaluate the latter, we may calculate the right-hand side of expression(9) more precisely, limiting the Bessel function expansion to two

terms of the series (which is sufficiently valid with $\mu \overline{A} < 3$). In this case

$$Y_e = \frac{2I_0\bar{v}L}{\omega U d\overline{A}} e^{j\frac{\pi}{2}} [X_0 e^{j\gamma} - X_0^3 (a\cos\gamma + jb\sin\gamma)] =$$

$$= \frac{2I_0\bar{v}L}{\omega U d\overline{A}} e^{j\left(\frac{\pi}{2}+\theta\right)} \left(X_0 - \frac{X_0^3}{2}\right)\left(\frac{1 - aX_0^2}{1 - \frac{X_0^2}{2}}\right) \times$$

$$\times \sqrt{1 + 2(a-b)\frac{1-(a+b)\frac{X_0^2}{2}}{(1-aX_0^2)^2} X_0^2 \sin^2\gamma}, \quad (13$$

where

$$X_0 \frac{\mu\overline{A}}{2} = \frac{X}{2};$$

$$a = \frac{\int_0^{\varphi_0} (p^2 + q^2) f(\varphi_0)\, d\varphi_0}{2\varphi_0^2 M^2(\varphi_0) \int_0^{\varphi_0} f(\varphi_0) d\varphi_0};$$

$$b = \frac{\int_0^{\varphi_0} (p^2 + q^2) \lambda(\varphi_0)\, d\varphi_0}{2\varphi_0^2 M^2(\varphi_0) \int_0^{\varphi_0} \lambda(\varphi_0) d\varphi_0}; \quad (13a$$

$$\operatorname{tg}\theta = \frac{1 - bX_0^2}{1 - aX_0^2} \operatorname{tg}\gamma. \quad (13b$$

Using the two terms of the series expansion of binomials and substituting $X_0 - \frac{X_0^3}{2} \simeq \simeq J_1(X)$, we may obtain from (13) and (13b) the expressions

$$Y_e = 2G_0 \frac{J_1(X)}{X} e^{-j(\delta+\pi)} [1 + (0{,}5 - a) X_0^2 - 0{,}5aX_0^4] \left\{1 + (a-b) \times \right.$$

$$\left. \times [1 - 0{,}5(a+b) X_0^2 - a(a+b) X_0^4] X_0^2 \sin^2 \frac{\varphi_0}{2}\right\}, \quad (1$$

$$\delta = -\frac{3\pi}{2} + \operatorname{arctg}\left\{[1 + (a-b) X_0^2 - abX_0^4] \operatorname{tg}\frac{\varphi_0}{2}\right\}, \quad (14a$$

where, after integration and substitution,

$$a = \frac{4\varphi_0^2 \sin\varphi_0 + \varphi_0(4 - \varphi_0^2)\cos\varphi_0 - \varphi_0\left(3 + \frac{2\varphi_0^2}{3}\right) + \left(\frac{\varphi_0^2}{2} - \frac{5}{4}\right)\sin 2\varphi_0 + \frac{3}{2}\varphi_0 \cos 2\varphi_0}{2\varphi_0^2 M^2(\varphi_0)[2\sin\varphi_0 - \varphi_0(1 + \cos\varphi_0)]},$$

(14b

$$b = \frac{\varphi_0(\varphi_0^2 - 6)\sin\varphi_0 + 4(\varphi_0^2 - 1)\cos\varphi_0 + \sin^2\varphi_0 + \left(\frac{\varphi_0^2}{2} - \frac{3}{4}\right)\cos 2\varphi_0 + \frac{19}{4} - \frac{3}{2}\varphi_0 \sin 2\varphi_0}{2\varphi_0^2 M^2(\varphi_0)[2(\cos\varphi_0 - 1) + \varphi_0 \sin\varphi_0]}.$$

From expressions (14) and (14a) it is evident that the more accurate initial relationship (11), the less quantities a and b differ from one another and from the value 0.5, and also the smaller the values of X_0 and $\sin\frac{\varphi 0}{2}$. Curves for the functions $a(\varphi_0)$ and $b(\varphi_0)$, plotted for the first and second zones of oscillation, are given in Figure 2. It follows from them that at that edge of each zone of oscillation, which corresponds small φ_0, the quantities a and b differ substantially from one another, whereupon a differs but little from 0.5. As follows from (14a), this will yield a definite error in determining the δ from formulas (11) but almost no error in determining Y_e from (11), since in expression (14) the factor of $\sin^2\frac{\varphi 0}{2}$ at this edge of the zone is close to zero. If at the center of the zone of oscillation

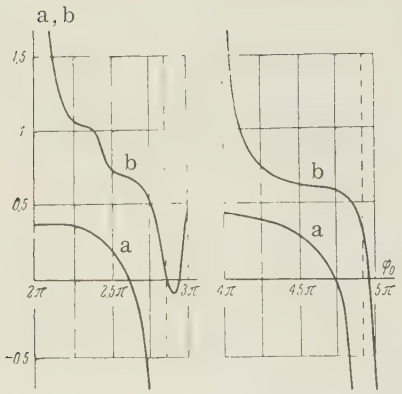


Figure 2.

$X < 2$, then the maximum content in (14a) of terms containing powers of X_0 will be less than 30%. At the edges of the zones of oscillation the error resulting from use of the approximate expressions in (11) decreases in conjunction with the decrease in $2X_0 = X$. Further analysis will be performed on the basis of the relationships in (11).

2. ANALYSIS OF OPERATION OF OSCILLATORS. COMPARISON WITH REFLEX KLYSTRON

Assuming that the oscillatory system is formed by the parallel connection of elements G_H, G, C and H (wherein G_H is the inductance of a matched load and G is determined by the losses in the resonator), we may obtain the following relationships characterizing operation of the oscillator:

$$P_н = \frac{G_н U^2}{2} = \frac{2G_н U_0^2 X^2 B_0^2}{M^2(\varphi_0)}, \quad \frac{X}{J_1(X)} = \frac{2G_0 \cos\delta}{G_н + G} = \frac{2I_0 F(\varphi_0)}{U_0 B^2 (G_н + G)}, \tag{15}$$

$$2\Delta\omega C = -(G_н + G)\operatorname{tg}\delta = (G_н + G)\operatorname{ctg}\frac{\varphi_0}{2}; \tag{16}$$

$$\Delta\omega = \omega - \omega_0; \quad \omega_0^2 = \frac{1}{LC}.$$

Using the expressions in (15) and the plot of function $X^2 = f\left[\frac{X}{J_1(X)}\right]$ [9], we may find the dependence of effective power (P_H) on angle φ_0 and thus plot the "zone of oscillation" of the device. Maximum power will occur at the center of the zone, where $\varphi_0 = 7.55$ for the first zone and $\varphi_0 = 13.9$ for the second zone. The frequencies generated at the centers of the zones (ω_c) will differ from the resonant frequency of the oscillatory system by the quantity

$$\Delta\omega_{c1} = \omega_{c1} - \omega_0 = \frac{0.675\omega_0}{Q_н}, \quad \Delta\omega_{c2} = \frac{0.606\omega_0}{Q_н}, \tag{17}$$

where $Q_H = \omega_0 C/(G_H + G)$ is the Q of the loaded resonator, wherein $\Delta\omega_{c1}$ refers to the first zone and $\Delta\omega_{c2}$ to the second.

Since the phase (δ) of electron admittance is equal to zero when $\varphi_0 = 3\pi, 7\pi, \ldots$, that is, outside the zone of oscillation, the diode oscillators cannot belong to the class of oscillators discussed in reference. [10]

It follows from (15) that the starting current of the oscillator is defined by the expression

$$I_{0п} = \frac{U_0 B^2 \omega_0 C}{Q_н F(\varphi_0)}. \tag{18}$$

Using (15), we may determine from the relationship dP_H/dG_H the optimum load conditions of the oscillator with given X:

$$J_0(X) = \frac{G}{G_0 \cos\delta} = \frac{GU_0 B^2}{I_0 F(\varphi_0)}, \quad J_2(X) = \frac{G_н}{G_0 \cos\delta} = \frac{G_н U_0 B^2}{I_0 F(\varphi_0)}. \tag{19}$$

It follows from (15) that efficiency of the oscillator is defined by the expression

$$\eta_н = \frac{P_н}{I_0 U_0} = \frac{2G_н F(\varphi_0) B_0^2 X^2}{G_0 M^2(\varphi_0) B^2 \cos\delta}. \tag{20}$$

Using (19) it is not difficult to reduce this relationship to the form

$$\eta_{\text{н}} = \frac{2I_0F^2(\varphi_0)B_0^2}{GU_0B^4M^2(\varphi_0)} X^2J_2(X)J_0(X). \tag{21}$$

With X = 1.84 (when $G_H = G$) the function $X^2j_2(X)J_0(X)$ has a maximum (reference [3]) of 0.338; hence, the maximum efficiency is defined by the expression

$$(\eta_{\text{н}})_m = 0{,}169\frac{I_0}{GU_0}\left[\frac{4F^2(\varphi_0)B_0^2}{B^4M^2(\varphi_0)}\right]. \tag{22}$$

It follows from (22) that with constant I_0/GU_0 in a monotron $(\eta_H)_m$ will be 1/32 of that in a diode with retarding field [since the monotron $B_0^2/B^4 = 1$ and not $32U_0^3/(U_0-U_2)^3$]. The maximum value of the quantity within brackets in expression (22) will be observed in the first zone of oscillation and is equal to 0.43 for a diode with retarding field. It may be shown (reference [3]) that in a reflex klystron the expression $(\eta_H)_m$ also has the form of (22), but instead of the factor within brackets there is $\beta^2\cos^2\delta$, where β is the coefficients of electron interaction. Since in a reflex klystron usually $\beta^2 \simeq 0.8$, it follows that $(\eta_H)_m$ in reflex klystrons is approximately 1.8 times greater than in diodes with retarding field.

From (15) and (16) we find that the total range of electronic tuning $(\Delta\omega_p)$ in diode oscillators is determined by the relationship

$$\frac{\Delta\omega_p}{\omega_0} = \frac{1}{2Q_{\text{н}}}\left(\operatorname{ctg}\frac{\varphi_{01}}{2} - \operatorname{ctg}\frac{\varphi_{02}}{2}\right) = \frac{\psi_p(K)}{Q_{\text{н}}}, \tag{23}$$

where $K = Q_HI_0 / U_0B^2\omega_0C$, where the function ψ_p is plotted on the condition that the quantities φ_{01} and φ_{02} are determined by means of the curves in Figure 1 and the relationship $KF(\varphi_0)=1$. Curves of the function $\psi_p(K)$ for the first and second modes are presented in Figure 3. For the second zone of oscillation the function ψ_p proceeds lower than for the first zone. Thus, the second and all subsequent zones of oscillation are less suitable with regard to $(\eta_H)_m$ and the range of electronic tuning in comparison with the first zone. From relationships (15) and (16) we find that the range of electronic tuning down to the half-power points $(\Delta\omega_{1/2})$ is defined by the expression

$$\frac{\Delta\omega_{1/2}}{\omega_0} = \frac{1}{2Q_{\text{н}}}\left(\operatorname{ctg}\frac{\varphi_{01}}{2} - \operatorname{ctg}\frac{\varphi_{02}}{2}\right) = \frac{\psi_{1/2}(X_m)}{Q_{\text{н}}}, \tag{24}$$

wherein φ_{01} and φ_{02} are found from the condition

$$F(\varphi_0) = \frac{J_1(X_m)F(\varphi_{0\text{п}})}{\sqrt{2}J_1(X_m/\sqrt{2})}, \tag{24a}$$

where X_m is the bunching parameter at the zone center when $\varphi_0 = \varphi_{0zc}$. The plot of function $\psi^1/_2(X)_m$ for the first zone is represented by the solid line in Figure 4. For a reflex klystron

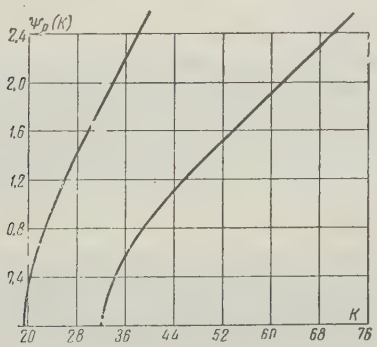


Figure 3

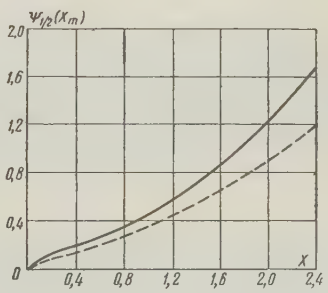


Figure 4

where X_m is the bunching parameter at the zone center when $\varphi_0 = \varphi_{0zc}$. The plot of function $\psi_{1/2}(X)_m$ for the first zone is represented by the solid line in Figure 4. For a reflex klystron the quantity $\Delta\omega_{1/2}$ may also be determined from expression (24) but the function $\psi_{1/2}(X_m)$ will in this case be found to be somewhat different (reference [4]). In Figure 4 the function $\psi_{1/2}(X$ for a reflex klystron is represented by the dashed line. It follows from Figure 4 that in diode oscillators in the first zone the range of electronic tuning to the half-power points is 30% greater than $\Delta\omega_{1/2}$ in reflex klystrons (it may be shown that in the second zone in diodes it is less than 15% greater). This is explained by shifting of the center of the zone of generation in diodes from the point $\delta = 0$ into the region of transit angles where $\operatorname{tg}\delta \simeq -1$. In this connection $\Delta\omega \sim \operatorname{tg}\delta$ will vary more effectively with a change in transit angle. This also explains a most important characteristic of diode oscillators, consisting in asymmetry in frequency change upon detuning from the zone center in different directions. Thus, with detuning in the direction of small values of φ_0 the frequency changes significantly more than with detuning in the other direction. This also explains the increase in sharpness of electronic tuning (at the zone center) in diode oscillators as defined by the relationship

$$-\left(\frac{d\omega}{d\varphi_0}\right)_{C1} = 1{,}42\,\frac{\omega_0}{2Q_н}, \tag{25}$$

in comparison with reflex klystrons, in which

$$-\left(\frac{d\omega}{d\varphi_0}\right)_C = \frac{\omega_0}{2Q_н}. \tag{25a}$$

If we consider the approximate expressions of (11) to be applicable, then, using the results obtained in the theory of single-cavity klystrons (reference [3]), we may show that the shortened equation for amplitude of oscillations within the diode has the form

$$\frac{1}{U}\frac{dU}{dt} = \frac{\omega_0}{2Q_н}\left[2\sigma\,\frac{J_1(X)}{X} - 1\right], \tag{26}$$

where $\sigma = G_0\cos\delta/(GH+G)$. Assuming in (26) that $2J_1(X) \simeq X - \frac{X^3}{8}$ and using the results of reference [3], we may derive the requirement of establishing oscillations in the diode:

$$U(t) = U_c\left[1 - \left(\frac{U_c^2}{U_т^2} - 1\right)e^{-\frac{\omega_0}{Q_н}(\sigma-1)t}\right]^{-1/2} \tag{27}$$

where U_c is the amplitude in steady-state operation, determined from (15); U_T is the initial surge in amplitude due to fluctuation phenomena. The set-up time of oscillations in the diode is then determined from the expression (reference [3])

$$t_y = \frac{Q_н}{\omega_0(\sigma-1)}\left\{\ln\left[Q_н(\sigma-1)\right] + \ln\left[\left(\frac{U_c}{U_т}\right)^2 - 1\right]\right\} \tag{28}$$

The influence of load on operation of diode oscillators is by nature similar to the influence of load on the operation of reflex klystrons and magnetrons. However, in distinction from the latter, the equal-frequency lines in the load characteristics of diode oscillators can never be parallel to the axis of conductances and for operation at the zone center they approach this axis at an angle of approximately 54°.

CONCLUSION

The above examination of operation of diode oscillators permits us to conclude that diode oscillators with a retarding field do not have significant advantages over reflex klystron oscillators with regard to the range of electronic tuning and maximum efficiency. The increase in sharpness of electronic tuning in the diode oscillator (1.42 times in comparison with the

sharpness of tuning of the reflex klystron) will lead to deterioration of stability of the generated frequency by 1.5 times.

The relationships derived above show that operation of diode oscillators within general lines is similar to operation of reflex klystrons and that all the examined types of diode oscillators may be investigated and calculated on one and the same principle. The output parameters of diodes with retarding field may be close to the values which are observed in reflex klystrons of similar rating. This is borne out by the data obtained on models of diode oscillators with retarding field (reference [7]).

REFERENCES

1. V.I. Kalinin, Generirovaniye detsimetrovykh i santimetrovykh voln, [Generation of decimeter and centimeter waves]; Svyaz'izdat, 1948.
2. V.M. Lopukhin, Vozbuzhdeniye elektromagnitnykh kolebaniy i voln elektronnymi potokami, [Excitation of electromagnetic oscillations and waves by electron streams]; GITTL, 1953.
3. S.D. Gvozdover, Teoriya elektronnykh priborov sverkhvysokikh chastot, [Theory of microwave electronic devices]; GITTL, 1956.
4. V.N. Shevchik, Osnovy elektroniki SVCh, [Principles of microwave electronics]; Sovetskoye Radio], 1959.
5. K. Wilmarth, J. Moll, Proc. IRE, 1952, 40, 7, 813.
6. J. Ebers, Proc, IRE, 1952, 40, 2, 138.
7. C.J. Carter, W.H. Cornett, M.O. Thurston, Vide, 1956, 12, 65, 281.
8. E. Grey, G.V. Matthews, Bessel functions and their applications in physics and mechanics, IL, 1949.
9. V.A. Malyshev, Radiotekhnika i Elektronika, 1960, 5, 10, 1603.
10. V.A. Malyshev, Izv. vuzov MVO SSSR (Radiofizika), 1959, 2, 3, 463.

Submitted to the editors 15 February 1960

CALCULATION OF PERIODIC ELECTROSTATIC FIELDS IN TRAVELING-WAVE TUBES WITH BIFILAR HELICES

A.L. Igritskiy

The report presents a method of calculating electrostatic fields in systems containing bifilar helices and designed for periodic electrostatic focusing of solid and hollow electron beams in traveling-wave tubes. Calculations are performed for the field within the bifilar helix, the field in the annular region bounded by the conducting cylinder and the bifilar helix, and the field in the annular region bounded by two bifilar helices having identical and different pitch.

INTRODUCTION

One of the most promising methods of electron beam focusing in TWT is that which employs periodic electrostatic fields. The advantages of this method are: the absence of any external focusing attachments; the method does not require adjustment of the position of the tube within the focusing arrangement, ensuring small size and weight. Bifilar helices are

usually used for the creation of periodic fields in TWT (in which the bifilar helix simultaneously performs the functions of focusing and retarding systems). In TWT with periodic electrostatic focusing both solid and hollow electron beams are used.

The subject of calculation of the periodic fields created by the bifilar helices has not been adequately discussed in the literature. Reference [1] gives a formula for potential distribution within the bifilar helix but does not indicate how it was derived. Reference [2] presents formulas for the potential distribution in a focusing model which consists of an assembly of cylinders. Instead of discussing the helix, reference [3] deals with the plane problem.

The present report presents a method of calculation of the electrostatic fields in systems containing bifilar helices and intended for the periodic electrostatic focusing of solid and hollow electron beams within TWT. The investigation revealed that use of the method of separation of Fourier variables in its usual form for calculation of the field of bifilar helices does not provide the desired result and that for solution of the problem it is necessary to transpose the independent variables.

The ensuing discussion includes: calculation of the field within the bifilar helix, calculation of the field in the annular region bounded by two bifilar helices having the same and different pitch.

1. CALCULATION OF THE ELECTROSTATIC FIELD WITHIN A BIFILAR HELIX

A bifilar helix is used for periodic electrostatic focusing of a solid beam of electrons (reference [1]). Figures 1a and 1b show a bifilar ribbon helix with radius r_1, ribbon width δ, and pitch L. To one ribbon of the bifilar helix there is applied a constant voltage V_0+V_f, and to the other a constant voltage V_0-V_f. If the spaced between ribbons of the helix is narrow, the electric field strength in this space may be considered constant and the potential distribution with $r=r_1$ may be taken to be approximately as shown in Figure 1c. The potential field within the helix must satisfy the Laplace equation, which in cylindrical coordinates (r,φ,z) has the form

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 V}{\partial \varphi^2}+\frac{\partial^2 V}{\partial z^2}=0. \tag{1}$$

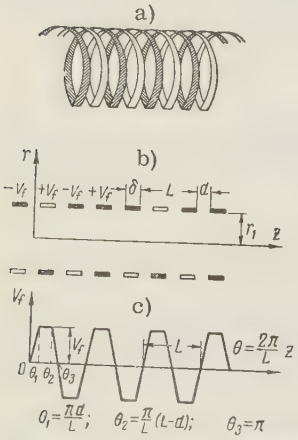


Figure 1. a, Bifilar ribbon helix; b, section with principal dimensions; c, approximate potential distribution in the plane $\varphi = 0$ with $r = r_1$.

Instead of determining the position of a point in cylindrical coordinates r,φ,z, let us give its position in the coordinates

$$R=\frac{2\pi}{L}r,\quad \theta=\frac{2\pi}{L}z-\varphi,\quad z. \tag{2}$$

Herein $\theta=\frac{2\pi}{L}z-\varphi=\text{const}$ is the geometric representation of a helical surface of constant pitch L, the coordinate R = const is a cylinder of radius R and z = const is the plane perpendicular to the helix axis. The intersection of these three surfaces also defines a point in space.

In the new variables the Laplace equation takes the form

$$\frac{\partial^2 V}{\partial R^2}+\frac{1}{R}\frac{\partial V}{\partial R}+\left(1+\frac{1}{R^2}\right)\frac{\partial^2 V}{\partial \theta^2}+\left(\frac{L}{2\pi}\right)^2\frac{\partial^2 V}{\partial z^2}+2\frac{2\pi}{L}\frac{\partial^2 V}{\partial z\partial\theta}=0. \tag{3}$$

In calculating the derivatives in (1) it is assumed that with the new variables V depends on φ through θ and depends on z directly and through θ.

In our case the electrodes between which the voltage is applied are bounded by helical lines. The position of these lines is wholly defined by giving two coordinates: R = const and θ = const. Thus the boundary conditions do not depend on the coordinate z. On this basis we shall be concerned with solution of the Laplace equation of $V(R,\theta)$ without dependence on z. For this case equation (3) may be rewritten in the form

$$\frac{\partial^2 V(R,\theta)}{\partial R^2}+\frac{1}{R}\frac{\partial V(R,\theta)}{\partial R}+\left(1+\frac{1}{R^2}\right)\frac{\partial^2 V(r,\theta)}{\partial \theta^2}=0. \tag{4}$$

Using Laplace equation (4), let us determine the potential distribution for the bifilar helix shown in Figure 1b. To one of the helices there is applied a focusing potential $-V_f$ and to the other $+V_f$. Exactly midway between the two helices there is a zero equipotential line which is also a helical line. Let us choose the position of the coordinate system (r, φ, z) relative to the bifilar helix so that the zero equipotential line $V = 0$ passes through the point with coordinates $r = r_1$, $\varphi = 0$ and $z = 0$. Then in the new coordinates, in accordance with (2), we may write the following boundary condition:

with
$$R = R_1 = \frac{2\pi}{L} r_1 \quad \theta = 0, \; V = 0. \tag{5}$$

Equation (4) is solved by the method of separation of Fourier variables (reference [4]). We shall seek the solution of this equation in the form of the product of two functions u and w, one of which depends only on R and the other only on θ:

$$V(R, \theta) = u(R)\, w(\theta). \tag{6}$$

Inserting this value of $V(R, \theta)$ into the Laplace Equation (4) and separating variables, we obtain

$$\frac{\frac{1}{u}\frac{d^2u}{dR^2} + \frac{1}{u}\frac{1}{R}\frac{du}{dR}}{1 + \frac{1}{R^2}} = -\frac{1}{w}\frac{d^2w}{d\theta^2}. \tag{7}$$

From (7) we obtain two ordinary differential equations for definition of w and u:

$$\frac{d^2w}{d\theta^2} + \nu^2 \omega = 0, \tag{8}$$

$$\frac{d^2u}{dR^2} + \frac{1}{R}\frac{du}{dR} - \nu^2\left(1 + \frac{1}{R^2}\right)u = 0, \tag{9}$$

where $\nu^2 = \text{const}$. Since the electric field varies periodically in the focusing arrangement, it is convenient to take $\nu^2 > 0$. Solving Equations (8) and (9) for functions w and u, we obtain the expressions

$$w = A_\nu \cos \nu\theta + B_\nu \sin \nu\theta, \tag{10}$$

$$u = C_\nu I_\nu(\nu R) + D_\nu K_\nu(\nu R). \tag{11}$$

Here $I_\nu(\nu R)$, $K_\nu(\nu K)$ are modified Bessel functions of the first and second kind, respectively, of order ν (reference [5]).

In accordance with (6), the potential is

$$V_\nu(R, \theta) = (A_\nu \cos \nu\theta + B_\nu \sin \nu\theta)\,[C_\nu I_\nu(\nu R) + D_\nu K_\nu(\nu R)]. \tag{12}$$

Since the potential is finite at the axis of symmetry of the system, the solution of (12) must not contain the function K_ν, which with $r = 0$ becomes infinite. Inserting into (12) boundary condition (5), we obtain $A_\nu = 0$. Thus, Equation (12) may be presented in the form

$$V_\nu(R, \theta) = B_\nu I_\nu(\nu R) \sin \nu\theta. \tag{13}$$

For coordinate θ the periodic field must have a period 2π. Proceeding from this, $\nu = n$ is an integer. Thus,

$$V_n(R, \theta) = B_n I_n(nR) \sin n\theta; \quad n = 1, 2, 3, \ldots. \tag{14}$$

The series obtained from these terms (with the coefficients B_n still subject to definition) must yield the required potential distribution

$$V(R, \theta) = \sum_{n=1}^{\infty} B_n I_n(nR) \sin n\theta. \tag{15}$$

With $R = R_1 = \frac{2\pi}{L} r_1$

$$V(R_1, \theta) = \sum_{n=1}^{\infty} B_n I_n(nR_1) \sin n\theta. \tag{16}$$

On the other hand, as is seen from Figure 1c, the potential distribution with $R = R_1$ in plane $\varphi = 0$ may, with a certain approximation, be represented by a trapezoidal curve. Thus, in order to determine the coefficients B_n it is sufficient to expand the trapezoidal curve of potential distribution into a series in terms of sines (reference [4]).

$$V(R_1, \theta) = \sum_{n=1}^{\infty} b_n \sin n\theta, \tag{17}$$

where

$$b_n = \frac{2}{\pi} \int_0^{\pi} V(R_1, \theta) \sin n\theta \, d\theta. \tag{18}$$

The integral for b_n is conveniently divided into a series of integrals, since the function $V(R_1, \theta)$ is represented as a broken line in the segment from 0 to π:

$$b_n = \frac{2}{\pi}\left[\int_0^{\frac{\pi d}{L}} \frac{V_f L}{\pi d}\theta \sin n\theta d\theta + \int_{\frac{\pi d}{L}}^{\frac{\pi}{L}(L-d)} V_f \sin n\theta d\theta + \int_{\frac{\pi}{L}(L-d)}^{\pi} \frac{V_f L}{\pi d}(\pi - \theta)\sin n\theta d\theta\right].$$

Evaluating these integrals and cancelling, we finally obtain

$$b_n = \frac{2}{\pi^2 n^2}\frac{V_f L}{d}\sin\left(\frac{n\pi d}{L}\right)(1 - \cos n\pi). \tag{19}$$

From comparison of (16) and (17) with consideration of (19) we have

$$B_n = \frac{2V_f}{I_n(nR_1)}\frac{1}{\sigma(n\pi)^2}\sin(n\pi\sigma)(1 - \cos n\pi), \tag{20}$$

where

$$\sigma = \frac{d}{L}. \tag{21}$$

Inserting the value of B_n from (20) into (15), we obtain the following expression for potential distribution:

$$V(R, \theta) = \sum_{n=1}^{\infty} \frac{2V_f}{\sigma(n\pi)^2}\frac{I_n(nR)}{I_n(nR_1)}(1 - \cos n\pi)\sin(n\pi\sigma)\sin n\theta.$$

Replacing the variables in accordance with (2) and also considering that, in addition to the focusing potential $\pm V_f$, there is applied to both ribbons of the helix the constant accelerating potential V_0 of the second anode of the electron gun. for the potential distribution within the bifilar helix we shall have the final expression

$$V(r, \varphi, z) = V_0 + \sum_n 4V_f \frac{\sin(n\pi\sigma)}{(n\pi)^2\sigma}\frac{I_n\left(\frac{2n\pi}{L}r\right)}{I_n\left(\frac{2n\pi}{L}r_1\right)} \times \sin\left(\frac{2n\pi}{L}z - n\varphi\right);\ n = 1, 3, 5, \ldots. \tag{22}$$

2. CALCULATION OF ELECTROSTATIC FIELD IN ANNULAR REGION BOUNDED BY CYLINDER AND BIFILAR HELIX

Investigations of periodic electrostatic focusing have shown that the best results are obtained with electron beams of annular cross section. To focus such a beam we may use the arrangement represented in Figure 2. To the cylinder 1, having a radius r_1, there is applied potential V_1. Voltages $V_0 + V_1$ and $V_0 - V_f$ are applied to the ribbons of the bifilar helix 2 of radius r_2. The voltage of the inner cylinder V_1 is slightly lower than the mean voltage V_0 of the bifilar helix. The annular beam of electrons passes into the region between the cylinder and the bifilar helix.

As before, the potential field in the annular region must satisfy the Laplace equation (1). Using the principle of superposition (reference [6]), let us divide the problem into two parts: first, let us calculate the field between two cylinders, the potentials of which are constant and equal to V_0 and V_1 (we neglect the bending of potential between the ribbons of the helix); second, let us calculate the field between the inner cylinder (the potential of which is equal to zero) and the bifilar helix (to the ribbons of which there is applied a voltage $\pm V_f$).

For the first problem the differential Laplace equation has the form

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dV}{dr}\right) = 0. \qquad (23)$$

The boundary conditions may be presented as:

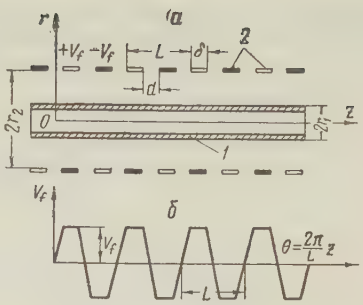


Figure 2. a, Diagram of the system with bifilar helix for periodic electrostatic focusing of hollow electron beam; b, potential distribution in plane $\varphi = 0$ with $r=r_2$.

$$\text{for } r = r_1 \quad V = V_1;$$

$$\text{for } r = r_2 \quad V = V_0. \qquad (24)$$

As the result of solution of Equation (23) with the boundary conditions in (24) we obtain for the potential the expression

$$V = V_0 - (V_0 - V_1)\frac{\ln\frac{r_2}{r}}{\ln\frac{r_2}{r_1}}. \qquad (25)$$

For the second problem the potential field in the annular region must satisfy the Laplace equation (4). The boundary conditions may be written in the following form:

$$\text{for } r = r_1 \quad V = 0. \qquad (26)$$

In the plane $\varphi = 0$ with $r = r_2$ the potential distribution must be represented by the trapezoidal curve shown in Figure 2b.

Solving Equation (4) by the method described in section 1, we obtain for the potential distribution equation (12). As before, midway between the two spirals there is the zero equipotential line of the periodic focusing field. Let us choose the state of the system of coordinates (r, φ, z) in such a manner (see Figure 2a) that with

$$R = R_2 = \frac{2\pi}{L} r_2 \quad \theta = 0, \quad V = 0. \qquad (27)$$

Inserting into (12) this boundary condition, we obtain $A_\nu = 0$. For coordinate θ the periodic field, as before, must have the period 2π. On this basis $\nu = n$ is an integer. Thus

$$V_n(R, \theta) = [C_n I_n(nR) + D_n K_n(nR)] \sin n\theta; \quad n = 1, 2, 3, \ldots \qquad (28)$$

The series obtained from the terms of (28) (with the coefficients C_n and D_n still subject to definition) must yield the required potential distribution

$$V(R, \theta) = \sum_{n=1}^{\infty} [C_n I_n(nR) + D_n K_n(nR)] \sin n\theta. \qquad (29)$$

Using boundary condition (26), we obtain the first equation for determining coefficients C_n and D_n:

$$C_n I_n(nR_1) + D_n K_n(nR_1) = 0. \qquad (30)$$

with

$$R = R_2 = \frac{2n\pi}{L} r_2$$

$$V(R_2, \theta) = \sum_{n=1}^{\infty} [C_n I_n(nR_2) + D_n K_n(nR_2)] \sin n\theta. \qquad (31)$$

On the other hand, as seen from Figure 2b, the potential distribution with $R = R_2$ in the plane $\varphi = 0$ may be represented as a trapezoidal curve. Thus, in order to determine coefficients C_n and D_n it is necessary to expand the potential distribution curve into a series in terms of sines

$$V(R_2, \theta) = \sum_{n=1}^{\infty} b_n \sin n\theta, \tag{32}$$

where

$$b_n = \frac{2}{\pi} \int_0^{\pi} V(R_2, \theta) \sin n\theta \, d\theta. \tag{33}$$

The integral for b_n is evaluated as in Section 1 and the quantity b_n will be defined by equation (19). Comparing (31) and (32), we obtain the second equation for determining coefficients C_n and D_n:

$$C_n I_n(nR_2) + D_n K_n(nR_2) = b_n. \tag{34}$$

By simultaneous solution of (30) and (34) we obtain

$$C = \frac{-b_n K_n(nR_1)}{I_n(nR_1) K_n(nR_2) - I_n(nR_2) K_n(nR_1)} \tag{35}$$

$$D = \frac{b_n I_n(nR_1)}{I_n(nR_1) K_n(nR_2) - I_n(nR_2) K_n(nR_1)} \tag{35'}$$

Inserting these values of C_n and D_n into Equation (29) and introducing instead of b_n its value from (19), with consideration of (2) and (21) we will have

$$V(r, \varphi, z) = 4V_f \sum_n' \frac{\sin(n\pi\sigma)}{(n\pi)^2\sigma} \frac{I_n\left(\frac{2n\pi}{L} r_1\right) K_n\left(\frac{2n\pi}{L} r\right) - K_n\left(\frac{2n\pi}{L} r_1\right) I_n\left(\frac{2n\pi}{L} r\right)}{I_n\left(\frac{2n\pi}{L} r_1\right) K_n\left(\frac{2n\pi}{L} r_2\right) - K_n\left(\frac{2n\pi}{L} r_1\right) I_n\left(\frac{2n\pi}{L} r_2\right)} \times$$

$$\times \sin\left(\frac{2n\pi}{L} z - n\varphi\right); \quad n = 1, 3, 5, \ldots . \tag{36}$$

The real potential at any point in the annular region between the cylinder and the bifilar helix is found as the sum of the potentials defined by formulas (25) and (36):

$$V(r, \varphi, z) = V_0 - (V_0 - V_1) \frac{\ln \frac{r_2}{r}}{\ln \frac{r_2}{r_1}} + 4V_f \sum_n' \frac{\sin(n\pi\sigma)}{(n\pi)^2\sigma} \times$$

$$\times \frac{I_n\left(\frac{2n\pi}{L} r_1\right) K_n\left(\frac{2n\pi}{L} r\right) - K_n\left(\frac{2n\pi}{L} r_1\right) I_n\left(\frac{2n\pi}{L} r\right)}{I_n\left(\frac{2n\pi}{L} r_1\right) K_n\left(\frac{2n\pi}{L} r_2\right) - K_n\left(\frac{2n\pi}{L} r_1\right) I_n\left(\frac{2n\pi}{L} r_2\right)} \sin\left(\frac{2n\pi}{L} z - n\varphi\right); \tag{37}$$

$$n = 1, 3, 5, \ldots .$$

3. CALCULATION OF ELECTROSTATIC FIELD IN ANNULAR REGION BOUNDED BY TWO BIFILAR HELICES OF IDENTICAL PITCH

Best results in focusing of hollow electron beams are obtained by the use of biperiodic focusing systems consisting of two bifilar helices located on the outside and inside of the hollow beam (see Figure 3). One of the bifilar helices (e.g., the outer) is used as the delay system. The inner bifilar helix is made of a highly resistive material and does not take part in the high-frequency interaction with the electron stream.

First, let us discuss the simple case of two bifilar helices having identical period L. Let the potential difference between the ribbons of the inner bifilar helix 1 be $2V_{f1}$ and

between the ribbons of the outer helix 2 let it be $2V_{f2}$. The mean potential of both bifilar helices is V_0.

The potential field in the space between the bifilar helices must satisfy Laplace equation (4). The boundary conditions will be: with $R = R_1 = \frac{2\pi}{L} r_1$ and $R = R_2 = \frac{2\pi}{L} r_2$ the potential must vary along the trapezoidal curves shown in Figures 3b and 3c, respectively.

Solving Equation (4) by the method described in Section 1, we obtain for the potential distribution Equation (12). For coordinate θ the periodic field, as before, must have period 2π. Accordingly, $\nu = n$ is an integer. We shall assume that the relative positions of the helices and the coordinate system are chosen as shown in Figure 3. Then the boundary conditions may be written in the following form:

$$\text{for } R = R_1 = \frac{2\pi}{L} r_1 \quad \theta = 0, \quad V = 0, \qquad (38)$$

$$\text{for } R = R_2 = \frac{2\pi}{L} r_2 \quad \theta = 0, \quad V = 0. \qquad (39)$$

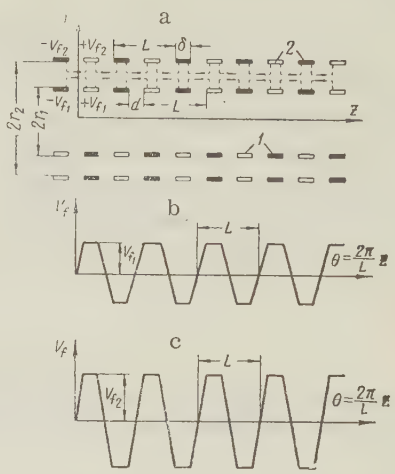


Figure 3. A, diagram of arrangement for biperiodic focusing of hollow electron beam; b, potential distribution in plane $\varphi = 0$ with $r = r_1$; c, potential distribution in plane $\varphi = 0$ with $r = r_2$.

Inserting boundary condition (38) into Equation (12) we have $A_n = 0$ and

$$V_n(R, \theta) = [C_n I_n(nR) + D_n K_n(nR)] \sin n\theta. \qquad (40)$$

The series obtained from the terms in (40) (with coefficients C_n and D_n still subject to definition) must provide the sought potential distribution

$$V(R, \theta) = \sum_{n=1}^{\infty} [C_n I_n(nR) + D_n K_n(nR)] \sin n\theta. \qquad (41)$$

При $R = R_1 = \frac{2\pi}{L} r_1$

$$V(R_1, \theta) = \sum_{n=1}^{\infty} [C_n I_n(nR_1) + D_n K_n(nR_1)] \sin n\theta. \qquad (42)$$

On the other hand, as seen from Figure 3b, the potential distribution in the plane $\varphi = 0$ with $R = R_1$ may be represented as a trapezoidal curve. Thus, in order to determine coefficients C_n and D_n it is necessary to expand the trapezoidal curve of potential distribution into the sine series

$$V(R_1, \theta) = \sum_{n-1}^{\infty} b_{n1} \sin n\theta, \qquad (43)$$

$$b_{n1} = \frac{2}{\pi} \int_0^{\pi} V(R_1, \theta) \sin n\theta \, d\theta. \qquad (44)$$

The integral for b_{n1} is evaluated as in Section 1. The quantity b_{n1} is defined by an equation similar to (19):

$$b_{n1} = \frac{2}{(n\pi)^2} \frac{V_{f_1} L}{d} \sin\left(\frac{n\pi d}{L}\right)(1 - \cos n\pi). \qquad (45)$$

By comparing Equations (42) and (43) we obtain the first equation for the determination of coefficients C_n and D_n:

$$C_n I_n(nR_1) + D_n K_n(nR_1) = b_{n1}. \tag{46}$$

In similar fashion, by examining the conditions at the second boundary, when $R = R_2 = \frac{2\pi}{L} r_2$, we obtain the second equation for determining coefficients C_n and D_n;

$$C_n I_n(nR_2) + D_n K_n(nR_2) = b_{n2}, \tag{47}$$

where b_{n2} is determined from equation (45) with substitution of V_{f2} for V_{f1}.

Solving simultaneously equations (46) and (47) we obtain the following equations for C_n and D_n:

$$C_n = \frac{b_{n1} K_n(nR_2) - b_{n2} K_n(nR_1)}{I_n(nR_1) K_n(nR_2) - I_n(nR_2) K_n(nR_1)}, \tag{48}$$

$$D_n = \frac{b_{n2} I_n(nR_1) - b_{n1} I_n(nR_2)}{I_n(nR_1) K_n(nR_2) - I_n(nR_2) K_n(nR_1)}. \tag{49}$$

Inserting these values of C_n and D_n into Equation (41) taking into account Equations (45) (2)1 and (21), we obtain the equation for potential distribution at any point between the helices

$$V(r,\varphi,z) = V_0 + \sum_n \frac{4\sin(n\pi\sigma)}{(n\pi)^2\sigma} \left\{ V_{f1} \frac{I_n\left(\frac{2n\pi}{L} r\right) K_n\left(\frac{2n\pi}{L} r_2\right) - I_n\left(\frac{2n\pi}{L} r_2\right) K_n\left(\frac{2n\pi}{L} r\right)}{I_n\left(\frac{2n\pi}{L} r_1\right) K_n\left(\frac{2n\pi}{L} r_2\right) - I_n\left(\frac{2n\pi}{L} r_2\right) K_n\left(\frac{2n\pi}{L} r_1\right)} + \right. \tag{50}$$

$$\left. - V_{f2} \frac{I_n\left(\frac{2n\pi}{L} r_1\right) K_n\left(\frac{2n\pi}{L} r\right) - I_n\left(\frac{2n\pi}{L} r\right) K_n\left(\frac{2n\pi}{L} r_1\right)}{I_n\left(\frac{2n\pi}{L} r_1\right) K_n\left(\frac{2n\pi}{L} r_2\right) - I_n\left(\frac{2n\pi}{L} r_2\right) K_n\left(\frac{2n\pi}{L} r_1\right)} \right\} \sin\left(\frac{2n\pi}{L} z - n\varphi\right);$$

$$n = 1, 3, 5, \ldots .$$

4. CALCULATION OF ELECTROSTATIC FIELD IN ANNULAR REGION BOUNDED BY TWO BIFILAR HELICES OF DIFFERENT PITCH

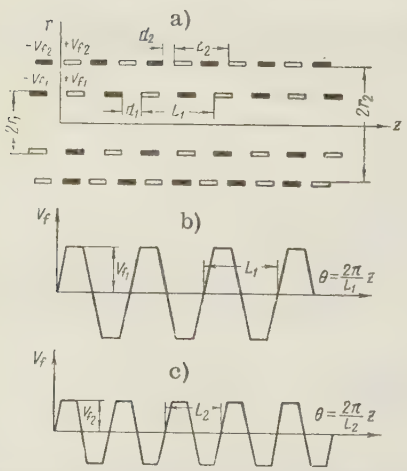


Figure 4. a) Diagram of arrangement for bi-periodic focusing of a hollow electron beam with the use of bifilar helices with different pitch; b) potential distribution in plane $\varphi=0$ with $r=r_1$; c) potential distribution in plane $\varphi=0$ with $r = r_2$.

Figure 4 represents a focusing system consisting of two bifilar helices with differe pitch. Let the potential difference between ribbons of the inner bifilar helix be $2V_{f1}$ an between ribbons of the outer helix $2V_{f2}$. Th mean potential of both bifilar helices is V_0. The potential field in the space between the bifilar helices must satisfy the Laplace equation (1). The boundary conditions will be as follows: with $r = r_1$ and $r = r_2$ the potential must vary along trapezoidal curve having periods L_1 and L_2 and shown in Figu 4b and 4c, respectively.

We shall solve this problem by use of t principle of superposition. Let us assume replacement of the inner bifilar helix with a conducting cylinder with zero potential and us calculate the field between this cylinder the outer bifilar helix, to the ribbons of whi there is applied a voltage $\pm V_{f2}$. Calculatio the field in this case is similar to the calcu tion performed in Section 2 and the potentia the space between the cylinder and the outer

bifilar helix may be determined from formula (36):

$$V_2(r, \varphi, z) = 4V_{f_2} \sum_n \frac{\sin(n\pi\sigma_2)}{(n\pi)^2\sigma_2} \frac{I_n\left(\frac{2n\pi}{L_2} r_1\right) K_n\left(\frac{2n\pi}{L_2} r\right) - K_n\left(\frac{2n\pi}{L_2} r_1\right) I_n\left(\frac{2n\pi}{L_2} r\right)}{I_n\left(\frac{2n\pi}{L_2} r_1\right) K_n\left(\frac{2n\pi}{L_2} r_2\right) - K_n\left(\frac{2n\pi}{L_2} r_1\right) I_n\left(\frac{2n\pi}{L_2} r_2\right)} \times$$

$$\times \sin\left(\frac{2n\pi}{L_2} z - n\varphi\right). \tag{51}$$

In the same fashion let us replace the outer bifilar helix with a conducting cylinder having zero potential and calculate the field between this cylinder and the inner bifilar helix, to the ribbons of which there is applied a voltage $\pm V_{f1}$. Then we obtain

$$V_1(r, \varphi, z) = 4V_{f1} \sum_n \frac{\sin(n\pi\sigma_1)}{(n\pi)^2\sigma_1} \frac{I_n\left(\frac{2n\pi}{L_1} r\right) K_n\left(\frac{2n\pi}{L_1} r_2\right) - I_n\left(\frac{2n\pi}{L_1} r_2\right) K_n\left(\frac{2n\pi}{L_1} r\right)}{I_n\left(\frac{2n\pi}{L_1} r_1\right) K_n\left(\frac{2n\pi}{L_1} r_2\right) - I_n\left(\frac{2n\pi}{L_1} r_2\right) K_n\left(\frac{2n\pi}{L_1} r_1\right)} \times \tag{52}$$

$$\times \sin\left(\frac{2n\pi}{L_1} z - n\varphi\right).$$

In Equations (51) and (52) $n = 1, 3, 5, \ldots$;

$$\sigma_1 = \frac{d_1}{L_1}; \quad \sigma_2 = \frac{d_2}{L_2}. \tag{53}$$

On the basis of the principle of superposition the total potential in the space between the helices will be

$$V(r, \varphi, z) = V_0 + V_1(r, \varphi, z) + V_2(r, \varphi, z). \tag{54}$$

REFERENCES

1. P.K. Tien, Focusing of a long cylindrical electron stream by means of periodic electrostatic fields, J. Appl. Phys, 1954, 25, 10, 1281.
2. K.K.N. Chang, Biperiodic electrostatic focusing for high-density electron beams, Proc IRE, 1957, 45, 11, 1522,
3. C.C. Jonson, Periodic electrostatic focusing of a hollow electron beam, IRE trans. Electron Devices, 1958, ED-5, 4, 233.
4. V.I. Smirnov, Kurs vysshey matematiki, [Course in higher mathematics], II, GTI, 1948.
5. E. Grey, G.V. Mathews, Bessel functions and their applications in physics and mechanics, IL, 1953.
6. V.A. Govorkov, Electric and magnetic fields, Svyaz'izdat, 1951.

Submitted to the editors 6 April 1960

INTERACTION OF MAGNETOSTATIC OSCILLATIONS IN FERRITE SPECIMEN DURING REGENERATION. PART I. INTERACTION OF SIMPLEST MODES

A. L. Mikaelyan and A. A. Vasil'yev

The report discusses the phenomena of microwave regeneration in a ferrite specimen. Interaction of the simplest modes of natural oscillation is investigated and the conditions required for their excitations are determined.

INTRODUCTION

Let us examine a ferrite sphere* magnetized in the z-axis direction by constant field of intensity H_0^e (Figure 1). Let the ferrite be acted upon by a strong magnetic field of frequenc ω_p. We shall consider that this field, known as the booster field has circular polarization wit clockwise rotation in the xy-plane. We shall designate the amplitude of this field by h_p^e. The magnetization excited in the ferrite will also consist of oscillations of clockwise rotation in the xy-plane. The frequency of these oscillations is ω_p and the relative amplitude is written in the form (reference [1])

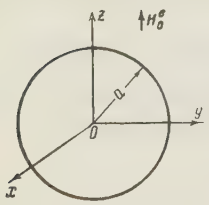


Figure 1

$$m_0 = \frac{M_p}{M} \frac{h_p^e}{\sqrt{(H_0^e - H_{res})^2 - (\Delta H_p)^2}}, \quad (1)$$

where $H_{res} = \omega_p/\gamma$ is the field corresponding to ferromagnetic resonance at frequency $\omega_p(\gamma = |e|/mc)$, ΔH_p is the half-width of the resonance curve of absorption, H_0^e and h_0^e are the external field M is the total magnetic moment, and M_p is the amplitude of the clockwise-polarized oscillation at frequency ω_p.

It is known (reference [2]) that under certain conditions in a ferrite specimen "disturbed" by a booster field it is possible to have self-excitation of oscillations at frequencies ω_1 and ω_2 satisfying the relationship

$$\omega_1 + \omega_2 = \omega_p. \quad (2)$$

This occurs because the action of the booster field causes a periodic change in the properties of the ferrite, which give rise to the possibility of excitation of oscillations at frequencies ω_1 and ω_2 coinciding with the resonant frequencies of natural oscillations of the ferrite specimen. The conditions under which the generation of oscillations occurs were first defined in reference [2] and then in reference [3]. A critique of these reports is given in the conclusion of the second part of this report.

*The phenomena discussed below may easily be applied to the case of a spheroidal ferrite specimen.

First we shall discuss the method of investigation of phenomena in a regenerated ferrite (Section1) and shall then illustrate it by examination of the simplest modes of oscillation (Section 2). In the second part of the report we discuss the interaction of more complex modes of oscillation, the possibility of which is denied in reference [3]. Finally, in the second part we derive a general formula for the regeneration threshold.

1. METHOD OF INVESTIGATION

In order to study the phenomena in a ferrite specimen it is necessary to examine the equations describing oscillatory processes in a disturbed ferrite at frequencies ω_1 and ω_2. In the case of small specimens, when

$$\operatorname{rot} \vec{H} = 0, \quad \vec{H} = \operatorname{grad} \Psi, \tag{3}$$

it is easy to derive these equations for potentials associated within the specimen with the magnetic moment by the following relationship:

$$\operatorname{div} \vec{B}_{1,2} = \operatorname{div} (\vec{H}_{1,2} + 4\pi \vec{M}_{1,2}) = 0, \quad \operatorname{div} 4\pi \vec{M}_{1,2} = -\operatorname{div} \operatorname{grad} \Psi_{1,2}, \tag{4}$$

where subscripts 1 and 2 refer to frequencies ω_1 and ω_2. Actually, by using the equation of motion

$$\frac{d\vec{M}}{dt} = -\gamma [\vec{M}\vec{H}] \tag{5}$$

and the condition of constant total magnetic moment, it is not difficult to determine the induction and magnetic moments at frequencies ω_1 and ω_2 due to the booster field (that is, in a disturbed ferrite). Then we obtain

$$\begin{aligned}
B_{x_1} &= \frac{\partial \Psi_1^i}{\partial x} + 4\pi M_{x_1} = \mu_1' \frac{\partial \Psi_1^i}{\partial x} - jk_1 \frac{\partial \Psi_1^i}{\partial y} - \tau_1 \frac{\partial \Psi_2^{i*}}{\partial z}, \\
B_{y_1} &= \frac{\partial \Psi_1^i}{\partial y} + 4\pi M_{y_1} = \mu_1 \frac{\partial \Psi_1^i}{\partial y} + jk_1 \frac{\partial \Psi_1^i}{\partial x} + j\tau_1 \frac{\partial \Psi_2^{i*}}{\partial z}, \\
B_{z_1} &= \frac{\partial \Psi_1^i}{\partial z} + 4\pi M_{z_1} = \frac{\partial \Psi_1^i}{\partial z} - \tau_2^* \left(\frac{\partial \Psi_2^{i*}}{\partial x} - i \frac{\partial \Psi_2^{i*}}{\partial y} \right),
\end{aligned} \tag{6}$$

where

$$\begin{aligned}
\mu_1 &= 1 + \frac{\omega_0 \omega_M}{\omega_0^2 - \omega_1^2}; \quad k_1 = -\frac{\omega_1 \omega_M}{\omega_0^2 - \omega_1^2}; \\
\tau_1 &= \frac{m_0}{2} (\mu_1 - k_1 - 1) = \frac{m_0}{2} \frac{\omega_M}{\omega_0 - \omega_1}; \\
\omega_M &= 4\pi\gamma M; \qquad \omega_0 = \gamma H_0^i.
\end{aligned} \tag{7}$$

In the derivation of (6) the only terms considered were those of the first order relative to $m_0 (m_0 \ll 1)$.

The induction at frequency ω_2 is described in the same manner as (6) with substitution of subscript 2 for subscript 1. After insertion of (6) into (4) we obtain the following equations for potentials within the ferrite at frequencies ω_1 and ω_2 (reference [3]):

$$\mu_1 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \Psi_1^i + \frac{\partial^2 \Psi_1^i}{\partial z^2} = (\tau_1 + \tau_2^*) \frac{\partial}{\partial z} \left(\frac{\partial}{\partial x} - j \frac{\partial}{\partial y} \right) \Psi_2^{i*}, \tag{8a}$$

$$\mu_2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \Psi_1^i + \frac{\partial^2 \Psi_1^i}{\partial z^2} = (\tau_1^* + \tau_2) \frac{\partial}{\partial z} \left(\frac{\partial}{\partial x} - j \frac{\partial}{\partial y} \right) \Psi_1^{i*}. \tag{8a}$$

In the absence of the booster field ($m_0 = 0$) the right-hand sides of Equations (8) vanish. In this case the system of (8), examined by Walker (reference [4]), describes the free "magnetostatic" oscillations at frequencies ω_1 and $\omega 2$. If the space within the sphere is described by spherical coordinates ξ, η, φ, then the solution of the homogeneous equations of (8) will be expressed in terms of adjoint Legendre polynomials

$$\psi_1 = \sum_n \sum_m A_{1,n,m}\, \psi_{1,n,m} = \sum_n \sum_m A_{1,n,m}\, P_n^{|m|}(\xi)\, P_n^{|m|}(\cos\theta)\, e^{jm\varphi}. \tag{9}$$

Here $A_{1,n,m}$ is the amplitude of internal oscillation at frequency ω_1. Subscripts n and m identify the structure of the given mode, wherein m may take either a positive (in the case of counterclockwise-polarized oscillations) or a negative sign (in the case of clockwise-polarized oscillation).

In the region outside the ferrite associated with the spherical system of coordinates

$$x = r\sin\theta\cos\varphi, \quad y = r\sin\theta\sin\varphi, \quad z = r\cos\theta, \tag{10}$$

the solution of (8) does not depend on the presence of the booster field (since here we always have $m_0 = 0$) and has the form

$$\Psi_1^e = \sum_n \sum_m D_{1,n,m}\, \frac{1}{r^{n+1}}\, P_n^{|m|}(\cos\theta)\, e^{jm\varphi}, \tag{11}$$

where $D_{1,n,m}$ is the amplitude of the external potential at frequency ω_1.

Detailed analysis of the solutions of the homogeneous equations of (8) is presented in references [4, 5] and, without dwelling on these matters, we shall turn to solution of a system of inhomogeneous partial differential equations.

We see that, as the result of action of the booster field, the two independent oscillations at frequencies ω_1 and ω_2 become interrelated. The greater the modulation of the ferrite properties (determined by the quantity m_0), the stronger this interrelation.

Solution of the system of (8) must be sought in the form of a sum

$$\begin{aligned} \Psi_1^i &= \psi_1 + \varphi_1, \\ \Psi_2^i &= \psi_2 + \varphi_2, \end{aligned} \tag{12}$$

where ψ_1 and ψ_2 are the solutions of the homogeneous Equations of (8) and φ_1 and φ_2 are the partial solutions of the mentioned inhomogeneous equations and result from the presence of the booster field. In addition, the amplitudes of potentials φ_1 and φ_2 are linearly associated with m_0. Limiting ourselves, as above, to the first order of m_0, we may consider that, for example, the right-hand side of Equation (8) depends only on potential ψ_2^*. Thus, φ_1 will be the partial solution of the equation

$$\mu_1\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\varphi_1 + \frac{\partial^2\varphi_1}{\partial z^2} = (\tau_1 + \tau_2^*)\,\frac{\partial}{\partial z}\left(\frac{\partial}{\partial x} - j\,\frac{\partial}{\partial y}\right)\psi_2^*. \tag{13}$$

Consequently, assigning a value of potential ψ_2^* (reference [5]), it is not difficult to find any partial solution of φ_1. These solutions are listed for a few cases in the following table

ψ_2	Right-hand side of Eq. (13)	φ_1
2,0,1	0	0
2,1̄,0	$2(\tau_1 + \tau_2^*)$	$(\tau_1 + \tau_2^*)\, z^2$
3,0,1	$-30(\tau_1 + \tau_2^*)(x - jy)$	$-15(\tau_1 + \tau_2^*)(x - jy)\, z^2$
3,1̄,0	$80(\tau_1 + \tau_2^*)\,\mu_2^*\, z$	$\frac{40}{3}(\tau_1 + \tau_2^*)\,\mu_2^*\, z^3$

Solution of inhomogeneous equation (8b) may be found by a similar method.

By virtue of the continuity of potentials at the ferrite-air boundary the value of φ_1 at the surface of the specimen determines the possible subscripts n and m for the sum of potentials ψ_1 corresponding to the solution of homogeneous equation (8a). In effect, by expanding φ_1 at the boundary into a Fourier series in terms of the entire system of orthogonal adjoint Legendre polynomials of $P_q^{|s|}(\cos\theta)$ we may write*

$$\varphi_{1\,|r=a} = \varphi_1(\theta)\, e^{js\varphi} = \sum_q a_q P_q^{|s|}(\cos\theta)\, e^{js\varphi}. \tag{14}$$

In addition, it must be noted that index s, determined according to (13) by potential ψ_2^*, is closely associated with index m of this potential in the following manner:

$$s = -(m+1). \tag{15}$$

The coefficients of a_q are defined by the expression (reference [6])

$$a_q = (2q+1)\,\frac{q-m+1}{q+m+1}\int_0^{\pi} \varphi_1(\theta)\, P_q^{|s|}(\cos\theta)\sin\theta\, d\theta. \tag{16}$$

It may be seen that (16) is other than zero if the following relationships are satisfied

$$q = n,\ n-2,\ n-4, \ldots, \tag{17}$$

where n is the index of potential ψ_2^*.

Analysis of (14) shows that potential ψ_1 within the ferrite is a sum of potentials in the form

$$\begin{gathered} A_{1n} P_n^{|m+1|}(\xi)\, P_n^{|m+1|}(\cos\theta)\, e^{-j(m+1)\varphi}, \\ A_{1(n-2)} P_{n-2}^{|m+1|}(\xi)\, P_{n-2}^{|m+1|}(\cos\theta)\, e^{-j(m+1)\varphi}. \end{gathered} \tag{18}$$

The corresponding sum outside the ferrite must also be taken.

Thus, as the result of interaction of potential ψ_2 and the booster field at frequency ω_1 there arises an entire spectrum of oscillations, the indexes of which are described in formulas (15) and (17). At frequency ω_2 there also will be excited a spectrum of oscillations with other indexes. Thus, excitation at frequency ω_1 of potential 5, 0, 1 will, with boosting at frequency ω_2, lead to excitation [in accordance with (15) and (17)] of the following spectrum of potentials: 5, $\bar{1}$, 0, 3, 1, 0, 1, $\bar{1}$, 0. In its turn, potential 5, $\bar{1}$, 0 excites the following potentials at frequency ω_1: 5, 0, 1, 3, 0, 1, 1, 0, 0.

It is important to note than in this case the higher potentials (e.g., 7, 0, 1 and 7, $\bar{1}$, 0) are not excited. Within the mentioned potential spectra there may appear those potentials which either are generally independent of the applied constant field (e.g. 1, 0, 0) or do not have resonant dependences (e.g., 2, 2, 0 et al.).

As seen from the above example, only those potentials which have the same first indexes prove to be interrelated.

This statement together with relationship (15) shows that in a regenerated ferrite specimen there may, for example, occur interaction of the following potentials: 2, 0, 1 - 2, $\bar{1}$, 0; 3, 0, 1 - 3, $\bar{1}$, 0; 4, 0, 1 - 4, $\bar{1}$, 0; 3, 1, 0 - 3, $\bar{2}$, 0; 4, 1, 0 - 4, $\bar{2}$, 0; 4, 2, 0 - 4, $\bar{3}$, 0, etc

Thus, the general rule for the indexes in interacting potentials has the form

$$\begin{aligned} n_2 &= n_1, \\ m_2 &= -(m_1+1). \end{aligned} \tag{19}$$

*As is known, the system of magnetostatic potentials is not complete. The requirement of completeness is necessary in the interest of convergence of series (14).

2. INTERACTION OF POTENTIALS 2, 0, 1 - 2, $\bar{1}$, 0

The simplest potentials interacting within the ferrite are 2, 0, 1 and 2, $\bar{1}$, 0. According to (17), in this case the spectrum at frequencies ω_1 and ω_2 consists only of the mentioned potentials. Let us assume that at frequency ω_1 there acts a potential 2, 0, 1 and at frequency ω_2 there acts a potential 2, $\bar{1}$, 0. Representing the potentials in the form (reference [5])

$$\begin{aligned} \psi_1 &= A_1\left(\frac{x^2+y^2}{2} - \mu_1 z^2\right) + C_1, \\ \psi_2 &= A_2 z\,(x - jy), \end{aligned} \tag{20}$$

where C_1 is a certain constant, we see that the right-hand side of Equation (8a) is equal to a constant and Equation (8b) is equal to zero. This means that $\varphi_2 = 0$. It is easily shown that

The choice of partial solution of φ_1 in any other form (e.g., in the form $A_2(\tau_1 + \tau_2)\frac{x^2+y^2}{2\mu_1}$ does not, of course, change the final results.

For the region within the ferrite we obtain

$$\begin{aligned} \Psi_1^i &= A_1\left(\frac{x^2+y^2}{2} - \mu_1 z^2\right) + A_2^*(\tau_1 + \tau_2^*)\, z_2, \\ \Psi_2^i &= A_2 z\,(x - jy). \end{aligned} \tag{21}$$

For the region outside the ferrite

$$\begin{aligned} \Psi_1^e &= D_1 \frac{1}{r^3} P_2^0(\cos\theta), \\ \Psi_2^e &= D_2 \frac{1}{r^3} P_2^1(\cos\theta)\, e^{-j\varphi}. \end{aligned} \tag{22}$$

From the condition of continuity of potentials at the ferrite-air boundary we find

$$\begin{aligned} D_1 &= \frac{2}{3} a^5 \left[A_2^*(\tau_1 + \tau_2^*) - \mu_1 A_1 - \frac{1}{2} A_1 \right], \\ D_2 &= \frac{1}{3} a^5 A_2, \end{aligned} \tag{23}$$

where a is the radius of the ferrite specimen. The induction outside the ferrite B_n^e, normal to the surface, is found as

$$\begin{aligned} B_{1n}^e|_{r=a} &= \left.\frac{\partial \Psi_1^e}{\partial r}\right|_{r=a} = -2a\left[A_2^*(\tau_1 + \tau_2^*) - A_1\left(\mu_1 + \frac{1}{2}\right)\right]\left[\frac{3}{2}\cos^2\theta - \frac{1}{2}\right], \\ B_{2n}^e|_{r=a} &= -3a A_2 \sin\theta\cos\theta\, e^{-j\varphi}. \end{aligned} \tag{24}$$

By means of (6) let us define the normal component of induction at the boundary within the ferrite*

$$\begin{aligned} B_{1n}^i|_{r=a} &= [\mu_1 A_1 - (\tau_1 + \tau_2^*) A_2^*]\, a - \\ &\quad - [\mu_1 A_1 - \tau_1 A_2^* + 2\mu_1 (\tau_1 + \tau_2^*) A_2^* + 2\tau_2^* A_2^*]\, a\cos^2\theta, \\ B_{2n}^i|_{r=a} &= \{[\mu_2 - k_2 + 1 - 2\tau_2(\tau_1^* + \tau_2)]\, A_2 + (2\mu_1^*\tau_2 - \tau_1^*) A_1^*\}\, a \sin\theta\cos\theta e^{-j\varphi}. \end{aligned} \tag{25}$$

*In expressions (25) second-order terms are not omitted. If they are disregarded, as was done in reference [3], the resulting potentials are not single-valued.

From the condition of continuity of induction we have

$$(4\mu_1 + 1) A_1 = 2(2\tau_1 + \tau_2) A_2,$$
$$[\mu_2 - k_2 + 4 - 2\tau_2(\tau_1^* + \tau_2)] A_2 = (\tau_1^* - 2\mu_1^*\tau_2) A_1^*. \tag{26}$$

Equations (26) are a system describing the interrelation of two potentials in the regenerated ferrite. Equating the determinant of system (26) to zero, we obtain

$$(4\mu_1 + 1)(\mu_2^* - k_2^* + 4) = 4\tau_1^2 + 4\tau_1\tau_2 + 2\tau_2^{*2} + 4\mu_1\tau_2^{*2}. \tag{27}$$

If the booster field is assumed equal to zero, the interrelation between oscillations disappears and system (26) decomposes into two independent equations

$$4\mu_1 + 1 = 0,$$
$$\mu_2^* - k_2^* + 4 = 0, \tag{28}$$

which describe the dependences of the resonant frequencies on the constant field and saturation magnetization of oscillations 2, 0, 1 and 2, 1, 0, respectively (see reference [5]).

Examination of Equation (27) shows that in the discussed system self-excitation of oscillations is possible at frequencies ω_1 and ω_2. In the presence of losses the characteristic Equation (28) becomes complex. In this case the natural frequencies of oscillation have the form

$$\dot{\omega}_1 = \omega_1 + j\omega''_{01}, \quad \dot{\omega}_2 = \omega_2 + j\omega''_{02},$$
$$\omega''_{01} = \gamma\Delta H_1; \quad \omega''_{02} = \gamma\Delta H_2; \quad \Delta H_1 \text{ и } \Delta H_2 \tag{29}$$

corresponding, respectively, to the half-widths of resonant lines of the first and second oscillations. Inserting (28) as the zero-th approximation in the right member of (27), we may determine the change in complex frequencies of (29) due to action of the booster field. In this case it develops that the increments of imaginary frequencies, caused by the booster field have signs opposite to those of $j\omega''_{01}$ and $j\omega''_{02}$ and may be greater in absolute value than the latter. In the case of total compensation of losses within the system the imaginary parts of frequencies as determined from (27) must be equal to zero, which corresponds to the absence of attenuation of oscillations at frequencies ω_1 and ω_2. This phenomenon will occur when

$$m_0 \geqslant 10 \frac{\sqrt{\Delta H_1 \Delta H_2}}{4\pi M} \frac{\sqrt{(\omega_0^2 + \omega_1^2)}}{3\omega_0 + \omega_1} = \frac{2 \mid \mu''_1 (\mu''_2 - k''_2)}{(\mu_1 - k_1 - 1) + 0{,}5 (\mu_2 - k_2 - 1)}, \tag{30}$$

where

$$\mu''_2 = \frac{\omega_M \omega''_{02} (\omega_0^2 + \omega_2^2)}{(\omega_0^2 - \omega_2^2)^2}; \quad k''_2 = -\frac{2\omega_0 \omega_M \omega_2 \omega''_{02}}{(\omega_0^2 - \omega_2^2)^2}; \tag{31}$$

the quantity μ''_1 is determined, as in the case of μ''_2, with replacement of subscript 2 by 1.

As we see, the threshold value of m_0 depends on the half-width of the line of each oscillation ΔH_1 and ΔH_2, the saturation magnetization and the choice of operating frequencies. Using equation (1), we find the following value of amplitude of the external magnetic field h_p^e at which the generation of oscillation commences*

$$h_p^e = 10 \sqrt{\Delta H_1 \Delta H_2} \frac{\sqrt{2\Omega_H^2 + \frac{4}{5}\Omega_H}}{3\Omega_H - \sqrt{\Omega_H^2 + \frac{4}{5}\Omega_H}} \times \sqrt{\sqrt{\left(\Omega_H^2 + \frac{4}{5}\Omega_H\right)} - \frac{2}{15}\Big)^2 + \left(\frac{\Delta H_p}{4\pi M}\right)^2} \tag{32}$$

*A similar expression with an error factor twice as great (see conclusion in part two of this report in reference [7]) is derived in reference [3].

where $\Omega_H = H_0^i / 4\pi M$.

If the frequency of the booster field is not resonant,

$$\left(\sqrt{\Omega_H^2 + \frac{4}{5}\Omega_H}\right)^2 \geqslant (\Delta H_p / 4\pi M)^2.$$

In this case for sufficiently large values of frequencies ω_1 and ω_2 ($\Omega_H >> 1$) the threshold formula is considerably simplified and takes the form

$$h_p^e \simeq (2.35 \cdot 10^{-3} H_0'' - 0.66) \sqrt{\Delta H_1 \Delta H_2}. \tag{33}$$

Figure 3 presents on an expanded scale the initial sector of the graph in Figure 2. Here we also find the resonance lines 2, 0, 1 and 2, $\bar{1}$, 0, the booster frequency and the resonance line 1. $\bar{1}$, 0 (all in relative frequency) and the relative quantity h_p^e. The point of intersection of line f_p / f_M and the line 1, $\bar{1}$, 0 (Ω_H = = 0.02) indicates that in this case the boost is applied at the resonant potential 1, $\bar{1}$, 0. This results in a sharp decrease in the threshold, which is now defined by the formula

$$h_p^e = 6.2 \frac{\Delta H_p}{4\pi M} \sqrt{\Delta H_1 \Delta H_2}. \tag{34}$$

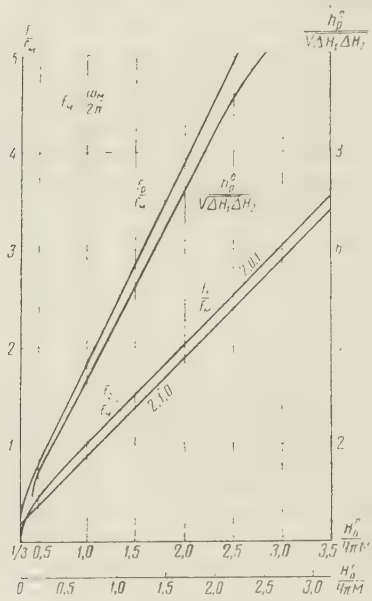


Figure 2

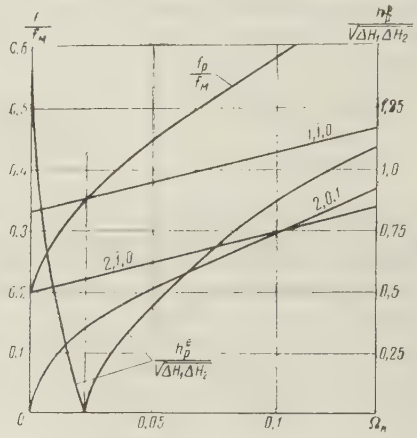


Figure 3

For a magnetization of $4\pi M = 1700$, gauss characteristic for yttrium ferrite, the threshold value of the external field is equal to

$$h_p^e = 3{,}65 \cdot 10^{-3} \Delta H_p \sqrt{\Delta H_1 \Delta H_2}. \tag{35}$$

At point $\Omega_H = 0.1$ the resonance frequencies of oscillations 2, 0, 1 and 2, $\bar{1}$, 0 are equal. This point corresponds to the "degenerate" condition — the excitation of oscillations of half frequency $\omega_{1,2} = 1/2\ \omega_p$).

In this case the threshold is

$$h_p^e = 0.88 \sqrt{\Delta H_1 \Delta H_2}. \tag{36}$$

We see that the critical value of the booster field in this case does not depend on saturation magnetization of the ferrite. Figure 4 gives curves for the dependence of the generation threshold of the discussed pair of potentials on the value of the external constant field for a magnetization of $4\pi M = 1700$ gauss. This figure shows the dependence of the resonance frequencies of potentials 2, 0, 1 and 2, $\bar{1}$, 0 on the value of H_p^e at the given value of saturation magnetization of the ferrite.

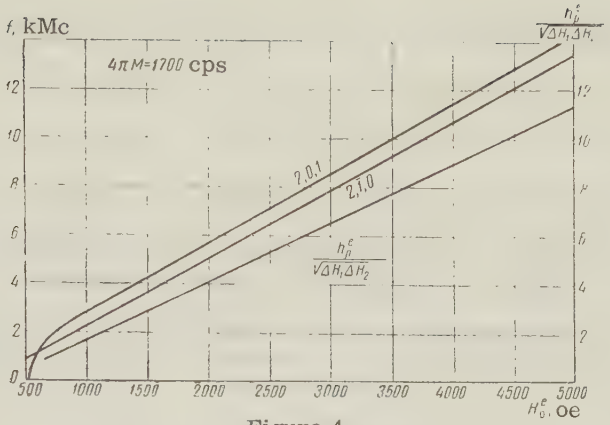


Figure 4.

REFERENCES

1. A. L. Mikaelyan, The problem of creation of microwave ferrite amplifiers, Radiotekhnika i elektronika, 1958, 3, 11, 1323
2. H. Suhl, Theory of the ferromagnetic microwave amplifier, J. Appl Phys., 1957, 28, 11, 1225.
3. Ya. A. Monosov, Toward a theory of nonlinear phenomena in ferrites at microwavelengths, Parts I-II, Radiotekhnika i elektronika, 1960, 5, 1-2, 59, 278.
4. L.R. Walker, Magnetostatic modes in ferromagnetic resonance, Phys. Rev. 1957, 105, 2, 39.
5. P. S. Fletcher, R.O. Bell, Ferromagnetic resonance modes in spheres, J. Appl. Phys., 1959, 30, 5, 687.
6. G. A. Stratton, Theory of electromagnetism GTTI, 1947, p. 355.
7. A. L. Mikaelyan, A.A. Vasil'yev, Interaction of magnetostatic oscillations in a regenerated ferrite specimen. Part II. Conditions of excitation of oscillations in the general case, Radiotekhnika i elektronika, 1961, 6, 5.

Submitted to the editors 21 April 1960

SOME PROPERTIES OF

A THICK-LAYER MULTIALKALINE PHOTOCATHODE

L. V. Kononchuk

This report discusses the manufacture of a multialkaline photocathode in the form of a thick layer.

The results of investigation of certain properties of a thick multialkaline photocathode are presented: the integral sensitivity, spectral characteristics and secondary-emission properties.

Also presented is a comparison of the investigated properties of a thick multialkaline photocathode with the properties of a semitransparent multialkaline photocathode and a few other technical photocathodes for the visible and infrared regions of the spectrum.

1. INTEGRAL AND SPECTRAL SENSITIVITY OF A THICK MULTIALKALINE PHOTOCATHODE

Until recently the multialkaline photocathode was made only in the form of a semitransparent layer (references [1,2,3,4,7]).

Sommer listed a maximum integral sensitivity for a semitransparent multialkaline layer of 200μa/lumen (T_{col} = 2870° K) (reference[2])

Mostovskiy and others list this same value of maximum sensitivity of a semitransparent multialkaline photocathode in individual specimens (T_{col} = 2854° K) (reference [4]).

Semitransparent multialkaline photocathodes in photocells and single-stage photoelectron multipliers developed by Soviet industry have sensitivities from 100 to 190 μa/ lumen (T_{col} = = 2854° K) (reference [7]).

It is interesting to note that despite the fact that the maximum sensitivities of semitransparent multialkaline photocathodes listed by various authors are approximately the same, the spectral curves which they give are different.

Thus, according to Sommer, the sensitivity of a semitransparent multialkaline photocathode in the short-wave region of the spectrum exceeds the sensitivity of an antimony-cesium photocathode in the same region (reference [2]). However, according to the data of other authors the sensitivity of a semitransparent multialkaline photocathode in the wavelength region from 300 to 580 microns does not exceed the sensitivity of an antimony-cesium photocathode (references [4,7]).

It is difficult to explain this discrepancy. We may only assume that the authors had differ photocathodes or we may search for possible errors in the method of comparison or of obtaining the spectral curves.

We obtained multialkaline photocathodes in the form of a thick layer with maximum integral sensitivity of 230 μ a/lumen (T_{col} = 2854° K) and spectral characteristics differing from the spectral characteristics of a multialkaline photocathode made in the form of a semitransparent layer.

The technology of preparing thick multialkaline photocathodes does not differ in principle from the technology of preparing semitransparent photocathodes as described in reference [8].

The change in color of a multialkaline photocathode upon treating a semitransparent layer and a solid layer of antimony with alkali metals is listed in Table 1.

In transmission emission and in reflection emission a semitransparent multialkaline (Sb-K-Na-Cs) photocathode has a yellow-brown or golden yellow color; a thick multialkaline photocathode has a blue-grey color in reflection emission and a yellow-brown or dark yellow color in transmission emission (to the extent that such emission is possible).

It is known that practical photocathodes — cesium-silver-oxide (Ag-O-Cs), antimony-cesium (Sb-Cs) and bismuth-silver-cesium (Bi-O- Ag- Cs) — made in the form of a thick layer have higher integral sensitivity than when these same cathodes are made in the semitransparent variant.

Table 1

Semitransparent photocathode			Solid photocathode		
Stage of preparation	Color		Stage of preparation	Color	
	"in transmission"	"in reflection"		"in transmission"	"in reflection"
Spraying of antimony layer to 40 percent decrease in transparency	dark color of metallic antimony	dark color of metallic antimony	Spraying of antimony layer to 40 percent decrease in transparency	bright color of metallic antimony	bright color of metallic antimony
Treating antimony layer with potassium	violet or lilac	violet or lilac	Treating layer with potassium	dark violet	greenish violet
Treating antimony-potassium layer with sodium	yellow-brown or golden yellow	yellow-brown or golden yellow	Treating antimony-potassium layer with sodium	yellow-brown	blue-gray
Treating antimony-potassium-sodium layer with cesium	no change in preceding	no change in preceding	Treating antimony-potassium-socium layer with cesium	no change in preceding	no change in preceding

The data for both types of these cathodes given in Table 2 are obtained from Sommer (reference[3]).

It must be noted that the semitransparent cesium-silver-oxide photocathode sometimes shows higher sensitivity than the thick variant. This deviation is explained by the fact that in this case the semitransparent cesium-silver-oxide photocathodes are made by a different technique than that used for thick photocathodes; for example, supplementary spraying of silver is required.

Table 2

Type of cathode	Integral sensitivity μa/lumen
Ag — O — Cs thick semitransparent	15—70 10—40
Sb — Cs thick semitransparent	30—110 15—60
Bi — Ag — O — Cs thick semitransparent	20—80 15—60

From the above remarks on the thick multialkiline photocathode it follows that it is also subject to the same rule of increased sensitivity as the above-mentioned photocathodes for the visible and infrared regions of the spectrum.

Figures 1 and 2 show the response curves of semitransparent and thick multialkaline photocathodes in relative units in comparison with other types of photocathodes; Figure 2 gives the response curves of photocathodes made in uviol envelopes.[1] For convenience of comparison these relative curves were recalculated as absolute curves in microamperes per

[1]The spectral characteristics of the cathodes were plotted in the illumination engineering laboratory under the direction of M. I. Epshteyn.

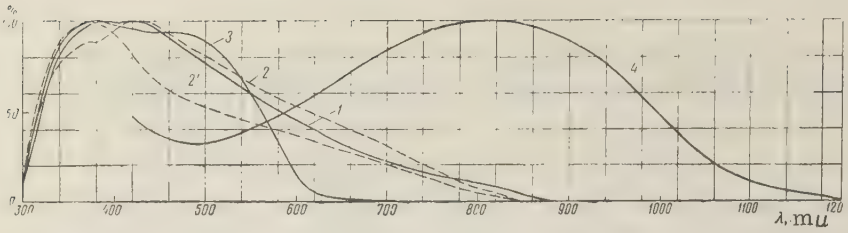


Figure 1. Curves for distribution of spectral sensitivity of photocells (in relative units): 1, photocell with reflection emission of thick multialkaline cathode, Σ_i = 210 μa/lumen; 2, photocell with transmission emission of semitransparent multialkaline cathode, Σ_1 = = 125 μa/lumen; 2', reflection emission, Σ_i = 79 μa/lumen; 3, photocell with reflection emission of solid antimony-cesium cathode, Σ_i = 112 μa/lumen; 4, photocell with reflection emission of solid cesium-silver-oxide cathode, Σ_i = 35 μa/lumen.

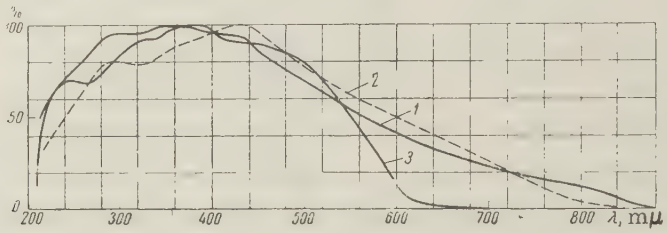


Figure 2. Curves for distribution of spectral sensitivity of photocells in uviol glass (in relative units); 1, photocell with reflection emission of thick multialkaline cathode, Σ_i = = 210 μa/lumen; 2, photocell with transmission emission of semitransparent multialkaline cathode, Σ_i = 98 μa/lumen; 3, photocell with reflection emission of thick antimony-cesium cathode, Σ_i = 100 μ2/lumen.

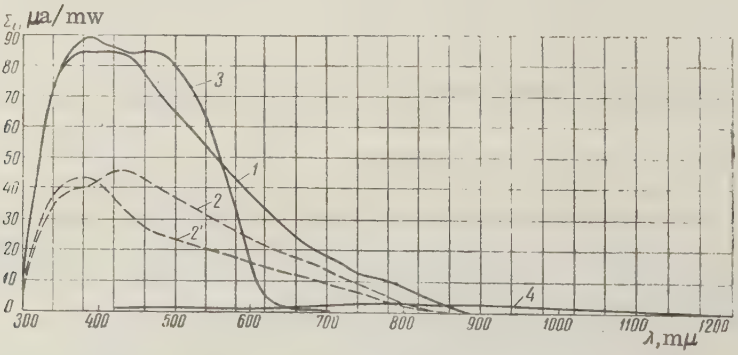


Figure 3. Curves for distribution of spectral sensitivity of photocells (in absolute units); 1, photocell with reflection emission of thick multialkaline cathode, Σ_i = 210 μa/ lumen; 2, photocell with transmission emission of semitransparent multialkaline cathode, Σ_i = = 125 μa/lumen; 2', reflection emission, Σ_i = 79 μa/lumen; 3, photocell with reflection emission of thick antimony cesium cathode, Σ_i = 112 μa/lumen; 4, photocell with reflection emission of thick cesium-silver-oxide cathode, Σ_i = 35 μa/lumen.

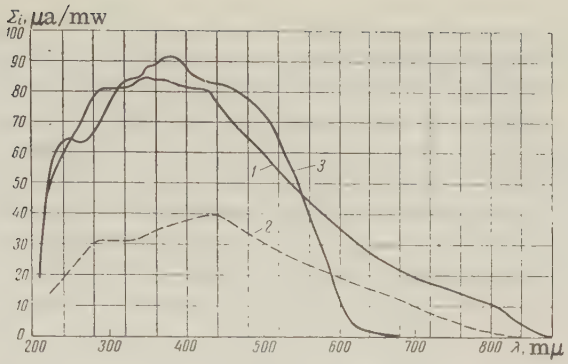


Figure 4. Curves for distribution of spectral sensitivity of photocells in uviol glass (in absolute units): 1, photocell with reflection emission of thick multialkaline cathode, Σ = 210 μa/ lumen; 2, photocell with transmission emission of semitransparent multialkaline cathode Σ_i = 98 μa/lumen; 3, photocell with reflection emission of thick antimony-cesium cathode, Σ_i = 100 μa/lumen.

milliwatt. The response curves of cathodes in absolute units are given in Figures 3 and 4. From an examination of these curves it is seen that, in addition to an overall increase in sensitivity for the thick multialkaline photocathode, there is a particularly significant increase in sensitivity in the short-wave region of the spectrum.

It is evident that the explanation of the increased sensitivity of thick layers and the change in their response curve in comparison with semitransparent layers must be sought in the change in spectral absorption of light in the working surface layer from which electron emission occurs.

Explanation of the increase in integral sensitivity in thick layers may also be connected with the fact that the conditions of light utilization in solid layers are more favorable than in semitransparent layers: by repeated reflection from the surface of crystals located beyond the working surface layer of the photocathode, light may return to this layer and cause an additional electron yield into the vacuum.

2. SECONDARY EMISSION PROPERTIES

Even in making single-stage multipliers with semitransparent multialkaline cathode and emitter it has been established that the semitransparent multialkaline layers possess secondary-emission properties not inferior to those of thick antimony-cesium cathodes (reference [7]).

The secondary-emission gain in the model developed averaged 5 - 11 with plate voltage of 220 v and emitter voltage of 170 - 200 v. At minimum cathode sensitivity of 100 μa/lumen the mean integral sensitivities of a single-stage photoelectron multiplier prove to be 500 - 100 μa/lumen. The maximum resulting sensitivity is 1300 μa/lumen. However, in the process of making such multipliers the secondary-emission properties of the multialkaline cathode were not compared with the secondary-emission properties of other materials.

In making single-stage multipliers with thick multialkaline cathode and emitter, a trial statistical comparison has been made between the secondary-emission properties of thick multialkaline layers and antimony-cesium layers

For such comparison single-stage multipliers were made with thick multialkaline and antimony-cesium cathodes and an emitter of one construction (the FEU-1 photoelectron multiplier construction was used). In making these and other multipliers, spraying of the antimony layer was achieved with identical thickness.

The first results of the secondary-emission properties of the mentioned layers permit the assumption that multialkaline layers are more effective emitters of secondary electrons than antimony-cesium layers (see Table 3).

From an examination of Table 3 it is seen that, first, the overall sensitivity of photomultipliers with thick multialkaline photocathode and emitter may be quite great and, second, the secondary-emission gain in a multialkaline layer is greater than in an antimony-cesium

Table 3

Type of photomultiplier	V_a = 220v			V_a = 500v		
	Integral sens. of cath. μa/lm	Overall sens. of multiplier μa/lm	Gain	Integral sens. of cath. μa/lm	Overall sens. of multiplier μa/lm	Gain
Single-stage multiplier with thick multialkaline cathode and emitter (average of 18)	131	1160	8.86	132	2160	15.75
Same (best result from above 18)	152	1700	11.15	157	3740	23.7
Single stage multiplier with thick antimony-cesium cathode and emitter (av. of 21)	97	692	7.16	96	1024	10.55
Same (best res. above 21)	110	1052	9.57	109	1495	13.75

layers, particularly with high anode voltages (i.e., with high primary-electron velocities).

Let us mention a phenomenon of considerable interest: on the first day of manufacture of multipliers with thick multialkaline cathode and emitter the gain at high voltages (in the given case, 500 v at anode) may reach extremely high values (25-37 per stage). Within 24 hours this value of gain drops sharply but still remains greater than in antimony-cesium layers and thereafter shows little or no change.

On the basis of the results of Spicer's investigation of the energy relationships in alkali-antimony photocathodes (reference [5]) we also assume in theory the high efficiency of the multialkaline photocathode in comparison with the antimony-cesium photocathode.

The energy structure of the multialkaline cathode differs from that of antimony-cesium chiefly in the width of the forbidden band: for the multialkaline photocathode, 1.0 ev; for the antimony-cesium photocathode, 1.6 ev; there is little difference in the width of the conduction band (in other words, the zone of electron affinity) in these cathode types (0.55 ev for the multialkaline and 0.45 for the antimony-cesium cathode) (reference [5]).

It is known that for dielectrics and intrinsic semiconductors the initial number of secondary electrons striking the conduction band may be defined as E_p/Q, where E_p is the energy o primary electrons and Q is the width of the forbidden band (reference [6]). In its most general form it is evident that this is also applicable for an extrinsic semiconductor.

In the given case, in a multialkaline photocathode the initial number of secondary electrons striking the conduction band must be greater than in an antimony-cesium photocathode, since the width of the forbidden band is vastly greater in the first of these photocathodes. Since in multialkaline and antimony-cesium photolayers the conduction bands are almost equal, it may be assumed that the principal factor influencing the increase in secondary-emission ratio is the difference in width of the forbidden bands.

It must be pointed out that precise quantitative investigations are required in order to make a final conclusion concerning the secondary-emission properties of the multialkaline photocathode.

It is evident that the solid multialkaline photocathode will find wide application in various photoelectric devices. Particularly promising is the use of such phctocathodes in those cases requiring sensitivity over a wide range of the spectrum (short-wave and long-wave), wherein the sensitivity in the short-wave region must not be lower than in the thick antimony-cesium photocathode.

REFERENCES

1. A. Sommer, Rev. Scient. Instrum., 1955, 26, 725.
2. A. Sommer, IRE Trans. Nucl. Sci. 1956, 3, 4, 8.
3. A. Sommer, Photocells (monograph), Methuin and Coe Publ., London, 2nd edition, 1951

4. A.A. Mostovskiy, O.V. Vorob'yeva, K. A. Mayskaya, Some properties of multialkaline photocathodes, Izv. AN SSR, Ser. fiz., 1958, 22, 5, 561.
5. W.E. Spicer, Photoemissive, photoconductive and optical absorption studies of alkali-antimony compounds, Phys. Rev., 1958, 112, 1, 114.
6. D. Wright, Semiconductors, transl. from English, IL, 1957.
7. T.N. Rabotnova, L.V. Kononchuk, Some data on the parameters and technological characteristics of photocells with multi-alkaline cathodes, Svetotekhnika, 1959, 9, 1.
8. L.V. Kononchuk, Technology of photocells and single-stage photoelectron multipliers with semitransparent multialkaline cathode, Radiotekhnika i elektronika, 1960, 5, 10, 1739.

Submitted to the editors 28 November 1960

MILLIKAN VACUUM CAPACITOR. PART I

M. Bakal, L.N. Dobretsov

The report discusses the physical fundamentals of the Millikan vacuum capacitor method and the structural diagram of an installation designed for study of the statistics of secondary electron emission.

INTRODUCTION

The chief characteristics of the phenomenon of secondary electron emission which have been studied in most works to date are the secondary emission ratio σ, the energy distribution of secondary electrons and the escape-angle distribution of secondary electrons.

The secondary emission ratio is the ratio of the current of secondary electrons to the current of primary electrons. This quantity is equal to the average number of secondary electrons liberated per primary electron.

It is of interest to determine the probabilities P_n of yield of one another number n of secondary electrons freed upon impact of a single primary electron. Distribution of these probabilities for the numbers n determines the characteristics of certain devices (e.g., the amplitude resolution and noise of a photoelectron multiplier, the efficiency of an electron multiplier as an electron counter). Knowledge of this distribution would permit determining certain parameters figuring in the theory of secondary electron emission.

The Millikan capacitor method may be used for measuring the number n. This is an extremely sensitive method of measuring a small change in charge of a body and was used by Millikan (reference [1]) in experiments demonstrating the atomic nature of electricity and measuring the electron charge; it was also used by A.F. Ioffe (reference [2]) in studying the elementary photoelectric effect. In one of the variants of this method a charged particle is held in equilibrium in the electric field of a parallel-plate capacitor. This is possible on the condition

$$qeE = mg,$$

where qe is the electric charge of the particle, e is the quantity of electron charge, E is the electric field strength of the parallel-plate capacitor, m is the mass of the particle and g is the acceleration of gravity.

In distinction from the experiments of Millikan and Ioffe, which were conducted in air

at atmospheric pressure, measurement of number n must be performed at a pressure below 10^{-3} - 10^{-4} mm Hg.

This circumstance leads to considerable complication in applying the Millikan capacitor method. Due to the absence of forces of internal friction in the vacuum, the particle has considerable velocity upon its introduction into the capacitor and within a short time acquires appreciable velocity during measurement of its charge. Hence, selection of the potential difference between the plates of the Millikan capacitor balancing the particle during its capture and during measurement of its charge, cannot be achieved directly by a man; it may be achieved only by means of an electronic system of automatic control.

Despite the complication of operating procedure in changing to a Millikan vacuum capacitor, it is also desirable that this be achieved because this method may be used for the solution of other physical problems ; for investigation of secondary emission under the influence of impacts of individual fast electrons and positrons; investigation of secondary electron emission under the influence of single ions; study of the charge of fine particles of powder in vacuum; etc.

Our development of the Millikan vacuum capacitor method was carried out on a model of the arrangement designed for measurement of the number of secondary electrons knocked out by a single primary electron.

1. PHYSICAL FUNDAMENTALS OF THE METHOD

In the Millikan vacuum capacitor the number of secondary electrons n freed by a single primary electron may be determined from measurement of the electric field intensities E_0 and E between the capacitor plates, which balance the particle (with known mass m) before and after it is struck by the primary electron, from the relation*

$$n = \frac{mg}{e}\frac{E - E_0}{LE_0} + 1. \tag{2}$$

In determining the mass of the particle in vacuum it is not possible to use the method used in Millikan's experiments, which is based on Stokes' law. In this connection the following method is proposed: in exposing the particle to electrons it is necessary to observe the cases of "adhesion" (that is, that absorption of the primary electron at which no secondary electron is dislodged.). Such cases will be observed particularly often in exposure of the particle to electrons with energy at which $\sigma < 1$.

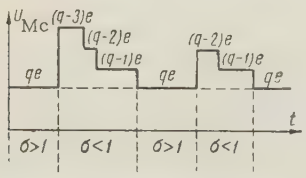


Figure 1. Schematic curve of change in voltage U_{mc} at Millikan capacitor in time with alternate exposure of particle to electrons with energy at which $\sigma > 1$ and exposure to electrons with energy at which $\sigma < 1$ (for negatively charged particle.

Cases of "adhesion" of single primary electrons are easily distinguished from cases of dislodging of any number $n > 1$ of secondary electrons, since they lead to voltage changes of opposite sign at that Millikan capacitor,

Considering that upon "adhesion" of a primary electron the change in particle charge is equal to the charge of one electron, the value of which is known, we may determine the mass of the particle from the relationship [easily determined from (2)]

$$m = \frac{e}{g}\frac{EE_0}{E_0 - E}. \tag{3}$$

In other words, we may omit measurement of the mass of the particle and even the field intensity. For this purpose, after each event of secondary electron emission with $n > 1$ it is necessary to change the energy of primary electrons to that at which $\sigma < 1$, and, upon reaching the initial voltage at the Millikan capacitor, to return to the condition $\sigma > 1$. The number of voltage "jumps" caused by adhesion of electrons is then equal to the number of secondary electrons dislodged from the surface of the particle during a secondary electron emission with $n > 1$ (Figure 1). Such

*All relations are derived for the case of a negatively charged particle. The advantage of operation with negatively charged particles is pointed out below.

operating conditions permit a large number of experiments with one particle.

The choice of particle sizes and operating range of field intensities of the Millikan capacitor must ensure the possibility of controlling voltage jumps corresponding both to adhesion of a single electron and to the yield of a sufficiently large number of electrons. The table lists the absolute values of jumps in field intensity of the Millikan capacitor with a change in particle charge per one electron charge for different values of particle mass and initial field intensity. From these data it follows that the above-mentioned requirements may be satisfied in operation with a particle, the mass of which is approximately equal to 10^{-10} g in the range of particle charge-to-mass ratios qe/m ≃ 300-25 CGSE. Such mass is possessed, for example, by a mercury droplet with radius 1.2μ or a spherical iron particle with radius 1.5μ. Retention within the Millikan capacitor of particles with qe/m ratios within the above range is possible with field intensities of the electric capacitor in the range from 100 to 12,00 v/cm.

It is evident that the installation does not directly record cases of dislodgement of only one secondary electron by a single primary electron. Determination of the number of these cases and the probabilities P_n is possible by two methods requiring additional measurements either of current density of primary electrons or of the secondary-emission ratio of the substance under investigation.

m, g	r_{Fe}, μ	r_{Hg}, μ	ΔE = ... for E_n = 2 kv/cm	ΔE = ... for E_n = 6 kv/cm	ΔE = ... for E_n = 12 kv/cm
10^{-10}	1.45	1.20	68	650	2900
$5 \cdot 10^{-10}$	2.48	2.06	14	120	492
10^{-9}	3.14	2.6	6	60	240

The first method consists in calculating the mean number of electrons striking the particle on the basis of measurement of the current density of primary electrons and the radius of the particle (the latter is easily calculated from (3) if the mass of the particle is measured. The difference between the mean number of primary electrons striking the particle over a certain interval of time and the number of electrons causing recordable changes in its charge over the same interval of time is equal to the number of primary electrons freed during the same time per secondary electron.

The second method, also discussed in reference [3], consists in the following: the ratios $N_n / N = R_n$ are compiled — the ratio of the number of cases of N_n of dislodgement of n secondary electrons to the total number N of recorded cases of change in charge of the particle over a certain interval of time. The ratios R_n are not equal to the probabilities P since the number N does not include the cases in which n = 1.

However, the ratio of values of R_n is equal to the ratio of probabilities P_n:

$$\frac{R_n}{R_e} = \frac{P_n}{P_e}. \tag{4}$$

In order to determine the absolute values of probabilities P_n it is necessary to use the relationships (see reference [4])

$$\sum_{n=0}^{\infty} P_n = 1, \tag{5}$$

$$\sum_{n=0}^{\infty} nP_n = \sigma. \tag{6}$$

In this method it is necessary to know the secondary emission ratio σ.

In studying the number of secondary electrons freed by a single primary electron it is essential that the particle be so irradiated that each recorded change in its charge is caused by the incidence of a single electron. This may be achieved by choosing a sufficiently small value of current density of primary electrons.

An important problem in the study of secondary electron emission in the Millikan vacuum capacitor is that of the determinacy of the energy of primary electrons reaching the surface of the particle. This energy (see Figure 2) is determined by the potential difference U_1 between the cathode of an electron gun and an anode connected with the upper plate of the capacitor and the surface of the particle. While potential difference U_1 may be easily varied, stabilized and measured, potential difference U_2 is unstable and is not precisely determined, since it

depends on a number of factors which vary in the course of the experiment: the potential differences between the capacitor plates, the vertical coordinate of the particle in the capacitor and the natural potential of the charged particle.

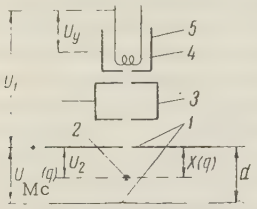


Figure 2. On the problem of determinacy of energy of primary electrons: 1, Millikan capacitor; 2, particle; 3, anode of electron gun; 4, cathode of electron gun; 5, control cylinder.

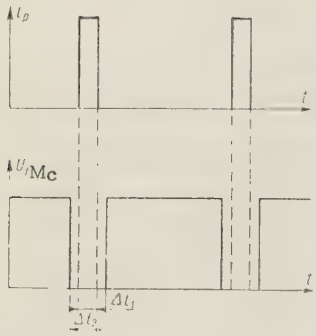


Figure 3. Time diagram of electron irradiation of particle: triggering of electron gun current i_p is achieved periodically at short intervals Δt_1 within periods Δt_2 when the voltage is removed from the Millikan capacitor.

One of the possible methods of eliminating the indeterminacy of U_2 consists in pulsed irradiation of the particle with electrons in the course of short periodic intervals of time Δt_1 during which the voltage from the plates of the Millikan capacitor is removed (see Figure 3)

Within the periods Δt_2 when the voltage is removed from the Millikan capacitor the particle behaves as a free falling body. If the periods Δt_2 are sufficiently small, the particle cannot be displaced over a significant distance Δx. For example, if $\Delta t_2 = 10^{-3}$ sec, $\Delta x = 5 \cdot 10^{-4}$ cm. This displacement is insignificant and does not result in loss of the particle.

This method does not eliminate the dependences of energy of the primary electrons on the natural potential of the particle. For particles with radius of approximately 1 micron and $qe / m \simeq 25 - 300$ CGSE the value of this potential is 0.06 - 0.0006 v and is small in comparison with the energy of the primary electrons. However, such value of surface potential may prevent escape of secondary electrons from the surface of a positively charged particle. In this connection it is advisable to conduct experiments with negatively charged particles.

The secondary emission ratio depends on the angle of incidence of the primary electrons. With arbitrary particle shape the electrons strike its surface at different angles and consideration of the angular relation becomes difficult. Such consideration is possible if the measurements are made with spherical particles. Operation with spherical particles is also desirable because the light flux scattered by the particle does not vary as the particle turns. As will be seen, this is of importance in operation of the automatic control system.

2. STRUCTURAL DIAGRAM OF THE INSTALLATION

The structural diagram of the installation may be divided into eight sections as shown in Figure 4.

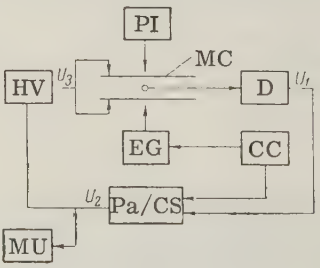


Figure 4. Structural diagram of installation: MC, Millikan capacitor, PI, particle injector; D, detector; Pa and CS, preamplifier and compensating section; HV, high-voltage amplifier; EG, electron gun; MU measuring unit; CC, control circuit.

In beginning the experiment a particle is introduced into the Millikan capacitor by particle injector PI. The automatic control system "captures" the particle and holds it in a "suspended" state at the center of the Millikan capacitor.

To the input of the preamplifier and compensating section (Pa and CS) there is applied a signal (voltage U_1) which is proportional to the vertical coordinate of the particle. Conversion of the coordinate into voltage is achieved by detector D. From the output of Pa and CS voltage U_2 reaches the input of high-voltage amplifier HV and measuring unit MU. From the output of HV voltage U is applied to the plates of the capacitor, whereby the automatic control system is closed. The changes in voltage U_2, proportional to changes in U_3, are controlled by measuring unit MU. Electron

bombardment of the particle is achieved by means of electron gun EG.

The last section shown in Figure 4, control circuit CC, controls the electron gun and amplifiers so that bombardment of the particle occurs in the absence of voltage between the capacitor plates. In addition, CC controls operation of the units of the installation at the beginning of the experiment, during capture of the particle, and performs certain other functions.

Description of the individual sections of the installation is given in reference [5].

REFERENCES

1. R. Millikan, Electrons (+ and -), protons, neutron and cosmic rays, ONTI, 1939.
2. A.F. Ioffe, Elementarnyy fotoelektricheskiy effekt, [Elementary photoeffect], Saint Petersburg, 1913.
3. R.E. Varrington, J.M. Anderson, Proc. Phys. Soc., London, 1958, 72, 467, 717.
4. M. Zeigler, Physica, 1936, 3, 1.
5. L.N. Afinogenov, M. Bakal, Ya. A. Fillippov, Millikan vacuum capacitor. Part II, Radiotekhnika i electronika, 1961, 6, 4, 658.

Submitted to the editors 28 November 1960

MILLIKAN VACUUM CAPACITOR. PART II

L. P. Afinogenov, M. Bakal and Yu. A. Filippov

The report describes the individual sections of an installation intended for study of the statistics of secondary electron emission.

INTRODUCTION

Part I (reference [1]) discussed the physical fundamentals of the Millikan vacuum capacitor method and described the basic layout of an installation intended for study of the statistics of secondary electron emission. This report describes the individual sections of this installation.

1. SECTIONS OF MILLIKAN VACUUM CAPACITOR INSTALLATION

1. Powder Particle Injector. For reliable capture of particles in the Millikan vacuum capacitor the method of introduction of the particles into the capacitor is of great importance. It is desirable that upon injection there be introduced one particle possessing low initial velocity and a charge-to-mass ratio lying within that interval to which the amplifier is tuned.

A method of injection similar to that used in reference [2] is most convenient: particles of metal powder contained in a metal reservoir (Figure 1) are drawn from the powder mass during the brief (10^{-3} sec) application of voltage between the reservoir and an anode near the surface of the powder. This anode is placed at the lower plate of the capacitor. A separating particle, reaching the Millikan capacitor through an aperture in the injector anode, is retarded in the gravitational field. The magnitude of the voltage pulse between the reservoir and the injector anode (500-1500 v) is so chosen that the particle loses all its velocity within the limits of the capacitor.

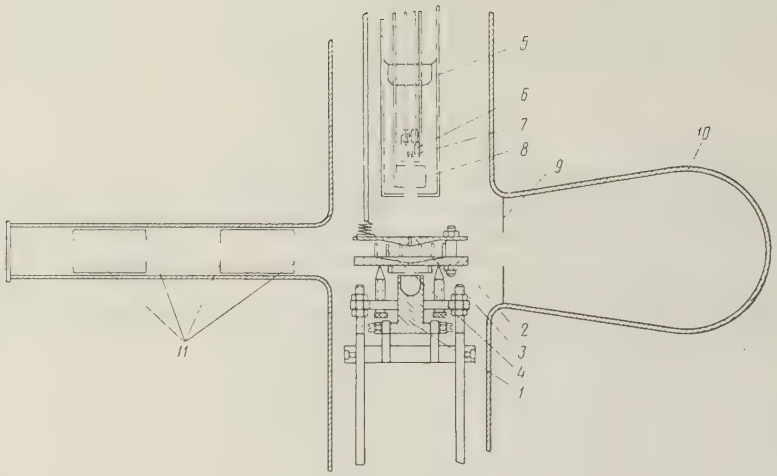


Figure 1. General view of vacuum device.
1, Metal reservoir for powder; 2, lower plate of Millikan capacitor; 3, upper plate of Millikan capacitor; 4, chielding rods; 5, electron gun shield; 6, control cylinder; 7, electron gun cathode; 8, anode; 9, diaphragm (D_0 in Figure 2); 10, envelope (T in Figure 2), 11 diaphragms (D_2 - D_5 in Figure 2).

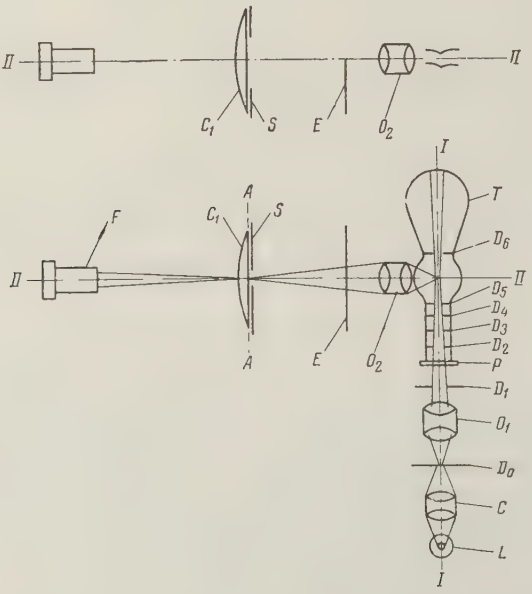


Figure 2, Diagram of illuminator and optical system.

Particles of carbonyl iron with spherical shape and diameter of 1 - 3 microns were used.

2. Detector. The detector, consisting of an illuminator and optical system with photomultiplier, provides at the output a voltage which is proportional to the vertical coordinate of the particle.

The automatic control system places two special requirements on the detector: (1) it is necessary to provide a low noise level at the output, for with the small particle dimensions the effective signal is extremely small. This is possible only with a sufficiently high ratio of light flux scattered by the particle, since the sensitivity of the photomultiplier is not identical with illumination of different portions of its cathode.

Figure 2 is a schematic diagram of the illuminator and the optic system.

The light source L is a DRSh-500 lamp; the lamp is powered by direct current from a storage battery. The positive column of the lamp is focused by lens C onto diaphragm D_0 with circular aperture of diameter 1.5 mm. This diaphragm is imaged by objective lens O_1 with focal length of 110 mm. in the plane II - II which passes through the axis of the Millikan capacitor. The light beam enters the vacuum tube through a plane-parallel ground glass window.

In order to decrease the light scattered by objective lens O_1 and window P and striking the capacitor plates the beam is limited by diaphragms D_1 - D_5. In leaving the capacitor the light beam passes through the aperture of diaphragm D_6 into chamber T of sufficiently large dimensions that only a small fraction of the beam entering it can return outward.

That portion of the light scattered by the particle in the vicinity of the capacitor axis and passing through the glass wall of the vacuum tube is focused by objective lens O_2 (a Yu-3 photo objective) onto the principal plane AA of planoconvex lens C_1, coinciding with its plane surface. Here is located a slot S which limits the incidence of parasitic light at the photomultiplier.

Between objective O_2 and lens C_1 a shield serves as an optical wedge and permits an approximately linear dependence of the light flux reaching lens C_1 on the vertical coordinate of the particle. Located at such a distance from O_2 that the light flux scattered by the particle is not focused, shield E suppresses part of this light flux, depending on the vertical coordinate of the particle. With downward displacement of the shield by 6 - 9 mm in the vicinity of the optical axis the light flux reaching the photomultiplier is varied from zero to maximum value.

Lens C_1 ensures immobility of the light beam (scattered by the particle) at the photomultiplier cathode upon displacement of the particle. The photomultiplier and lens C_1 are so located that the image of the input aperture of objective O_2 is in the plane of the photomultiplier cathode.

A specially chosen FEU-19 photomultiplier with low dark current was used.

3. Millikan Capacitor. Particle Stabilization in the Horizontal Plane. The automatic control system does not act on the particle position in the horizontal plane. Hence, any horizontal component of velocity (for example, that attained by the particle during introduction into the capacitor) may cause displacement of the particle from the capacitor axis and lead to its loss.

In the described device stabilization of the particle in the horizontal plane is achieved by horizontal components of the electric field of the capacitor directed toward the axis. The plates of the capacitor represent two conical surfaces with vertices pointing downward (2 and 3, Figure 1). Diameter of the plates is 5.4 cm and the distance between them 0.8 cm. With such configuration of the electric field the presence of a small horizontal component in particle velocity (0.5 cm/sec) will lead only to oscillation of the particle in the horizontal plane.

Stability of the particle in the horizontal plane may be ensured only with axially symmetrical distribution of potential around the capacitor axis. Such distribution may be disturbed by leads passing near the capacitor and by axial asymmetry of the walls of the glass envelope. Hence, voltage is applied to the capacitor plates through leads sealed into opposite stems of the vacuum device. The influence of axial asymmetry of the walls of the device is eliminated by metal rods 4 located with axial symmetry on the lower plate of the capacitor.

In the center of the lower plate there is an 0.8 mm orifice through which the iron particles are introduced into the capacitor. The electron beam enters the capacitor through an orifice in the center of the upper plate.

4. Electron Gun. Irradiation of the particle by electrons of different energies is

achieved by an electron gun of simple design within a shield (Figure 1).

It was pointed out in reference [1] that statistical material on one and the same particle may be obtained only by alternating its irradiation by electrons with energy at which the secondary emission ratio $\sigma > 1$ with irradiation by electrons with energy corresponding to $\sigma < 1$. In working with iron particles the energy of primary electrons corresponding to $\sigma < 1$ was chosen in the energy region above the "second unit" on the curve for dependence of σ on the energy of primary electrons.

5. Structural Scheme of Electrical Portion of the Installation. The electrical sections of the installation must ensure fulfilment of the following basic functions:

1. Capture of the particle introduced into the capacitor on the condition that its charge-to-mass ratio and initial velocity lie within permissible but sufficiently wide limits.

2. Stabilization of the charged particle within the capacitor by means of a system for automatic control of the voltage at the capacitor plates. In particular, the automatic control system must hold the particle in the event of a sudden change in charge within the limits likely to be encountered in the course of the experiment.

3. Irradiation by electrons of specified energy within a somewhat wide range.

4. Measurement of changes in particle charge due to impacts of single electrons.

In order to achieve functions 1 and 2 it is desirable to increase the operating speed of the control system. However, the subject of high-speed action must be considered together with the subject of noise, for the methods of combating noise signals reduce to affecting the frequency response of the system and, consequently, its operating speed.

The light signal from the particle, striking the photomultiplier, is extremely small in view of the small dimensions of the particle and the limited brilliance of the light source. Hence, the noise signal at the photomultiplier output (due, in part to the amplifier noise) and the inherent (dark) noise of the photomultiplier (due, in part, to parasitic scattered light as well as to noise arising at the photomultiplier output upon action of the light signal from the particle) may prove comparable with the effective signal.

In connection with these problems two systems of amplification and conversion of the signal from the photomultiplier output were compared: a system with d-c amplifiers and a system with a-c amplifiers. This comparison led to the following conclusions.

1. In a system with a-c amplifiers the influence of amplifier noise and inherent (dark) noise of the photomultiplier may be reduced. However, this noise is considerably less than the noise caused by parasitic scattered light and the light signal from the particle. The use of light modulation does not permit a significant reduction in the last two noise components, since they also modulate the carrier frequency along with the fundamental signal.

Consequently, with regard to anti-noise characteristics in the present case the a-c system can hardly offer significant advantages.

2. A shortcoming of the system with d-c amplifiers is the inherent zero drift. It may be decreased by the use of special circuits, stabilization of supply voltages and temperature. Moreover, in the presence of a captured particle the control system proves to be closed and the influence of drift is sharply decreased.

3. The system with a-c amplifiers is more complex than the system with d-c amplifiers not only due to the presence of auxiliary units but also due to the complexity of differentiation a-c signals.

On the basis of these considerations a system with d-c amplifiers was used in the described model.

Function 3 is achieved by the use of the pulse method of particle irradiation by electrons (see reference [1]): the electron gun is usually "blocked" by selection of the control electrode potential and is "opened" only periodically at short intervals of time. In these intervals irradiation of the particle by electrons occurs and during the irradiation the voltage is removed from the Millikan capacitor.

The structural diagram of the electrical portion of the installation is shown in Figure 3. The particle is illuminated by a constant light flux. The light scattered by the particle is converted by the detector D into a constant electrical signal which is proportional to the vertical coordinate of the particle. The signal from the detector output is amplified by d-c amplifiers A_1, A_2, A_3 and high-voltage amplifier HV, from the output of which the signal is applied to the Millikan capacitor.

There are compensating circuits: a differentiator DIF developing a derived signal required for stability of the control system, and an integrator INT, decreasing the static error of the system.

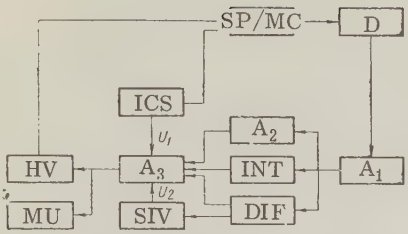


Figure 3. Structural diagram of Millikan vacuum capacitor installation; SP in MC, spherical particle in Millikan capacitor: D, detector; A_1, A_2, A_3, d-c amplifiers; HV high-voltage amplifier; SIV, section for initial voltage; ICS irradiation control section; MU, measuring unit; INT and DIF, integrator and differentiator.

The signals from A_2, DIF and INT are summed by amplifier A_3. In addition, two control voltages are applied to A_3: U_1 from the irradiation control section (ICS) and U_2 from the section providing the initial voltage (SIV).

The irradiation control section (ICS) controls the electron gun and amplifier, achieving irradiation of the particle in the time interval when the voltage is removed from the Millikan capacitor. It contains a pulse generator, feeding pulses to the control electrode of the electron gun, and an amplifier. Pulse voltage U_1 reaching the in put of A_3 ensures removal of the high voltage from the Millikan capacitor at the moment of electron irradiation of the particle.

Voltage U_2 from SIV acts only during capture of the particle. When the particles are injected into the Millikan capacitor they possess initial velocities. In order to capture a particle it is necessary that it is first retarded in the capacitor. Upon injection of a particle through the orifice in the lower plate of the capacitor it is retarded in the field of gravity. There is no voltage at the capacitor; the control system is cut off until the moment when the velocity of the particle becomes close to zero, for the force created by the control system acts upward against the force of gravity. SIV creates the blanking voltage and determines the moment of its removal.

The voltage at the output of A_3 is proportional to the voltage at the Milikan capacitor (during operation of HV in the linear portion of the characteristic). To the output of A_3 there is connected a measuring unit MU, consisting of a cathode follower and loop oscilloscope with recording of the indications on photographic paper. The MU permits examination of the change in particle charge during the course of the experiment.

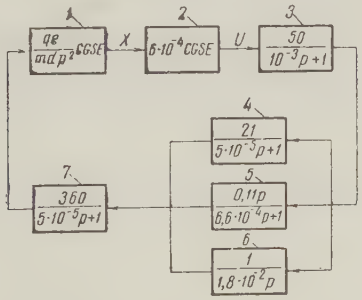


Figure 4. Design circuit of automatic control system: 1, particle in Millikan capacitor, qe, particle charge; m, particle mass, d, distance between capacitor plates; 2, detector: 3, first and second stages of amplifier (A_1 in Figure 2): 4, third and fourth stages of amplifier (A_2 and A_3 in Figure 2); 5, differentiator and fourth stage of amplifier; 6, integrator and fourth stage of amplifier; 7, high-voltage amplifier.

The design circuit shown in Figure 4 was adopted for dynamic calculation of the automatic control system. This figure shows the transfer functions of the individual links. With the parameters shown in this circuit the system is stable.

2. TESTING THE INSTALLATION

Tests of the installation showed that the automatic control system is capable of capturing spherical iron particles with diameter of 2-microns and retaining them in a stable state for a long period of time.

The automatic control system holds these particles stable even during the brief periodic removal of voltage from the Millikan capacitor

Irradiation of negatively charged particles by electrons with energy at which $\sigma > 1$ leads to an increase in voltage at the capacitor. Irradiation of these particles by electrons with energy at which $\sigma < 1$ leads to a decrease in voltage at the capacitor, which is to be expected. Incidence of a single electron at the particle causes a change in voltage at the capacitor by jumps. This is illustrated by oscillograms of the voltage at the capacitor (Figure 5) in the absence of irradiation and in the presence of irradiation of the particle by electrons when $\sigma > 1$ and $\sigma < 1$. The mass of the particle, calculated from relationship (3) (reference [1])

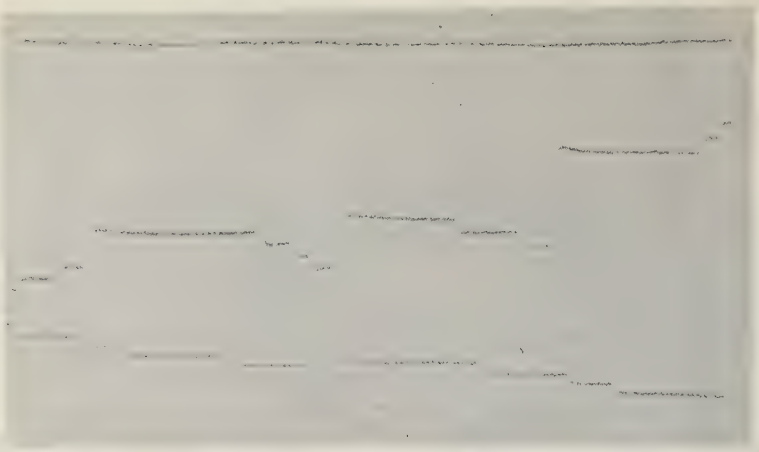


Figure 5. Oscillograms of voltage at Millikan capacitor; a, in the absence of irradiation, $U_{MC} = 7500$ v; b, irradiation of particle with $\sigma > 1$; c, irradiation of particle with $\sigma < 1$.

is $1.26 \cdot 10^{-9}$ g. A change in particle charge by one electron charge corresponds to a voltage jump of approximately 75 v. Downward jumps correspond to adhesion of primary electrons and upward jumps correspond to liberation of $n > 1$ secondary electrons. The relatively sloping leading edges of the jumps are explained by the time constant of the measuring unit used.

The authors express their sincere thanks to L.N. Dobretsov, under whose guidance the work was performed, and also to D.V. Shapot and V.A. Ivanov, who took part in designing the automatic control system.

REFERENCES

1. M. Bakal, L.N. Dobretsov, Radiotekhnika i elektronika, 1961, 6, 4, 653.
2. R.F. Wuerkler, H. Shelton, R.V. Langmuir, J. Appl. Phys., 1959, 30, 3, 342.

Submitted to the editors 28 November 1960

BRIEF COMMUNICATIONS

SHANNON SCHEME FOR GAUSSIAN INFORMATION WITH UNIFORM SPECTRUM AND FOR A CHANNEL WITH FLUCTUATION NOISE

B.S. Tsybakov

Contemporary demonstrations of the Shannon theorem (reference [1]) are based on the widely-used Shannon concept of random coding of long segments of information. The final, mathematically rigorous proofs of the theorem in the case of continuous transmission takes such an unwieldy form (for example, see reference [2]) that its perception becomes extremely difficult. In principle, there evidently exist simpler demonstrations for direct confirmation of the theorem for certain classes of channels and information sources which are based on previously known (found or guessed) methods of optimal encoding and decoding.

In the present remarks we wish to call attention to one simple demonstration of the Shannon theorem in the special case of "white" Gaussian information and channel. We will show that linear methods of encoding and decoding in the discussed case are optimal. For demonstration we have used the well known and developed theory of optimal linear filtering.

Let the output of the information source ξ (t) be a stationary, random Gaussian process with uniform spectrum in the band of frequencies W (spectral density). Let us designate its spectrum as $f_\xi(\omega)$:

$$f_\xi(\omega) = \frac{Q}{W}, \tag{1}$$

where Q is the amount of information at the output of the source.

For transmission of the information let there be a channel at the output of which the signal η'(t) is formed from the signal η (t) at the input in accordance with the relation

$$\eta'(t) = \eta(t) + \zeta(t), \tag{2}$$

where ζ (t) is independent of η (t) and is a stationary, random Guassian process with uniform spectrum in the same frequency band W as ξ (t). Let us designate its spectrum by $f_\xi(\omega)$:

$$f_\zeta(\omega) = \frac{P_n}{W}, \tag{3}$$

where P_n is the additive noise power in the channel.

We shall assume that the signal at the channel input η (t) possesses a given finite power P_s.

The problem consists in determining whether (direct theorem) or not (inverse theorem) there exist such methods of encoding and decoding as will permit transmission with a given accuracy of reproduction of the information at the output.

We shall assume that the permissible deviation of the decoded information at the output ξ'(t) from the initial information ξ (t) will be the mean-square error σ^2:

$$\sigma^2 = M[\xi(t) - \xi'(t - T)]^2,$$

where M is the averaging and T is the delay of the information at the output relative to the information at the input.

It is known (reference [1]) that for the described situation methods of encoding exist if

$$C = W \log\left(1 + \frac{P_s}{P_n}\right) \geqslant W \log \frac{Q}{\sigma^2} = R, \tag{4}$$

and do not exist when R > C.

In performing the operations of encoding and decoding, generally speaking, it is necessary to achieve a certain delay T in the information to be transmitted. If the values of R and C are such that R < C, that is,

$$\sigma^2 > \frac{QP_n}{P_s + P_n} = \sigma_0^2, \tag{5}$$

then the required delay T is a finite quantity. With a decrease in σ^2 the required value of T, generally speaking, increases and when $\sigma^2 \to \sigma_0^2$ it approaches infinity.

Let us examine linear filtering in the capacity of the encoding and decoding operations. Let the encoding consist in linear filtering

$$\eta_1(t) = \int_{-T_1}^{\infty} K(\tau)\, \xi(t - \tau)\, d\tau, \tag{6}$$

where $K(\tau)$ is the transient response of the filter, and in subsequent delay of the filter output by time T_1

$$\eta(t) = \eta_1(t - T_1).$$

The encoding operation may be achieved by means of a single, physically realizable, linear filter with transient response $K_1(\tau) = K(\tau - T_1)$.

Let us assume that the decoding is achieved by a linear filter

$$\xi'(t) = \int_{-T_1}^{\infty} M(\tau)\, \eta'(t - \tau)\, d\tau \tag{7}$$

with delay in T_1. The total delay in transmission arising due to encoding and decoding is $T = 2T_1$.

In order to evaluate the optimum transmission, let us examine filters with infinite delay* We shall show that under the given conditions they will be best for the encoding and decoding operations as described by Shannon.

In fact, according to the theory of optimal linear filtering the least mean-square error $\tilde{\sigma}_0^2$ under the given conditions has the form (see, for example, references [3, 4])

$$\tilde{\sigma}_0^2 = \int \frac{f_\xi(\omega)\, f_\zeta(\omega)\, d\omega}{|k(\omega)|^2 f_\xi(\omega) + f_\zeta(\omega)} = \int \frac{QP_n\, d\omega}{W(|k(\omega)|^2 Q + P_n)}, \tag{8}$$

where $|k(\omega)|^2$ is the square of the absolute value of the spectral characteristic of the encoding filter; it is chosen from the conditions

$$\int |k(\omega)|^2 f_\xi(\omega)\, d\omega = P_s, \tag{9}$$

$$|k(\omega)|^2 = \left[A \sqrt{\frac{P_n Q}{W^2}} - \frac{P_n}{W}\right] \frac{W}{Q}. \tag{10}$$

In formulas (8) and (9) integration is performed for bandwidth W; in formula (10) the quantity A is an undefined constant**.

On the basis of (9) and (10) we have

$$|k_0(\omega)|^2 = \frac{P_s}{Q}, \tag{11}$$

after which we obtain for $\tilde{\sigma}_0^2$

$$\tilde{\sigma}_0^2 = \frac{QP_n}{P_c + P_n}. \tag{12}$$

*By choosing sufficiently large T_1 we may approximate filtering with infinite delay to any desired accuracy.

**The expression for $\tilde{\sigma}_0^2$ and relations (9) and (10) were first given by Costas in reference [3]. More recently these results were obtained independently by Sinay (reference [4]) in whose wo the calculations relating to their derivation are described in greater detail.

Comparing this expression with (5), we see that the discussed methods of encoding and decoding are the best methods in terms of the Shannon concept. In this way (converting from (12) to (4) and replacing the equality sign with an inequality sign) we have proved Shannon's direct theorem for a white Gaussian information source and channel.

That the linear methods of encoding and decoding are optimal among all possible methods is explained by the fact that in the given case the source itself develops the optimal signals for the given channel. The problem of the filters consists only in the change in power of the corresponding processes. It may be shown that even in the apparently simple case of a Gaussian source with nonuniform spectrum the optimal encoding with the above channel is essentially nonlinear. The search for the best methods of encoding and decoding requires development of the theory of optimal nonlinear filtering.

In conclusion let us note that the mean-square error $\tilde{\sigma}^2_0$ will be determined from formula (12) also in the case where processes with uniform spectrum ζ (t) and ξ (t) are non-Gaussian. However, in this case condition (4) is replaced by another condition and $\tilde{\sigma}_0^2$ will not be minimum in all possible encoding and decoding methods.

REFERENCES

1. C.E. Shannon, A mathematical theory of communication, Bell, System Techn. J. part I 1948, 27, 379-423; part II, 1948, 27, 623-656.
2. R.L. Dobrushin, General formulation of the basic Shannon theorem in information theory, Uspekhi matem. nauk, 1959, 14, 6, (90), 3-104.
3. J.P. Costas, Coding with linear systems, proc, IRE, 1952, 40, 9, 1101.
4. Ya. G. Sinay, Least error and the best method of transmission of stationary information with linear encoding and decoding in the case of Gaussian communications channels, Problemy peredachi informatsii, [Problems in information transmission]; Izd. AN SSSR, 1959, 2, 40.

Submitted to the editors 26 November 1960

TOWARD A RECEIVING ANTENNA THEORY

B. Ye. Kinber

In reference [1] there was derived a formula for the power received by an antenna upon the incidence of a nonplanar wave. In deriving this formula it was assumed that the dissipated power is much less than the received power, and, consequently, that the incident wave approximately satisfies the boundary conditions at the walls of the antenna.

Nevertheless, in cases of practical interest (e.g., incidence of the wave at the antenna from the direction of the side lobes of the pattern, calculation of the mutual influence of closely spaced antennas) the power dissipated by the antenna considerably exceeds the power passing through the antenna system and the initial assumptions of reference [1] are clearly invalid.

In the present remarks it is shown that the formula for received power derived in reference [1] is rigorous and applicable given any relation between the received and dissipated fields.

Let a field $\vec{E}^n$, $\vec{H}^n$ be incident at a receiving antenna matched with a long feed in which

only one mode may be propagated. The $\vec{E}^n$, $\vec{H}^n$ wave is the principal field and does not satisfy the boundary conditions at the walls of the antenna. The virtual source of the $\vec{E}^n$, $\vec{H}^n$ field is some other transmitting antenna, which we shall consider as a system of given external currents. Thereby we avoid analysis of multiple diffractions of the field at the receiving and transmitting antennas.

At the receiving antenna part of the $\vec{E}^n$,$\vec{H}^n$ field enters the channel and part of it is dissipated (diffracted). The resulting field, which satisfies the boundary conditions, will be designated by $\vec{E}$,$\vec{H}$.

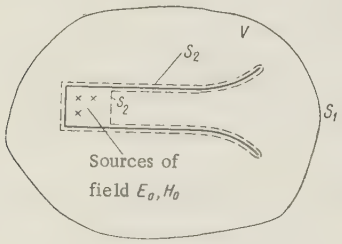
According to the equivalence theorem (reference [2]), the system of external currents, being virtual sources of the principal field, may be replaced by a combination of electric j_e and magnetic j_m currents at closed surface S_1 surrounding these sources with a surface density

$$\vec{j}_e = [\vec{n}\vec{H}^n], \tag{1}$$

$$\vec{j}_m = -[\vec{n}\vec{E}^n], \tag{2}$$

where $\vec{n}$ is the exterior normal to S_1.

Let V represent the volume bounded by S_1 and containing the receiving antenna. According to reference [2] the sources defined by expressions (1) and (2) excite field $\vec{E}'$, $\vec{H}'$ which within V coincides with the total field $\vec{E}$, $\vec{H}$.



Let us now examine the $\vec{E}_0$, $\vec{H}_0$ field of a receiving antenna during transmission operation; we shall consider this field to be known. The $\vec{E}_0$, $\vec{H}_0$ power flux will be chosen equal to unity. In the antenna channel the $\vec{E}_0$, $\vec{H}_0$ field is a wave traveling toward the "throat" of the antenna. Outside the antenna $\vec{E}_0$, $\vec{H}_0$ is the radiation field — a divergent spherical wave: $\vec{E}' = \vec{E}$, $\vec{H}' = \vec{H}$ (3)

Let us write the Lorentz lemma for the $\vec{E}_0$, $\vec{H}_0$ and $\vec{E}'$,$\vec{H}'$ fields. We shall integrate over the entire infinite space bounded externally by the surface of an infinitely remote sphere and internally by surface S_2 coinciding with the antenna surface and by section s_2 in the feed (see Figure 2).

Since the integral of an infinitely remote sphere and surface S_2 is equal to zero, the Lorentz lemma has the form

$$\int_{s_2} (\{[\vec{E}'\vec{H}_0] - [\vec{E}_0\vec{H}']\}, \vec{n})\, ds = \int_{s_1} \{(\vec{j}_e\vec{E}_0) - (\vec{j}_m\vec{H}_0)\}\, ds. \tag{4}$$

Substituting for $\vec{j}_e$ and $\vec{j}_m$ their expressions from (1) and (2), replacing $\vec{E}'$,$\vec{H}'$ with $\vec{E}$,$\vec{H}$, in accordance with (3) and remembering (from reference [3]) that

$$\int_{s_2} (\{[EH_0] - [E_0H]\}, \vec{n})\, ds = C_{-0}N_{-0}, \tag{5}$$

where

$$N_{-0} = \int_{s_2} (\{[\vec{E}_0\vec{H}_0^*] + [\vec{E}_0^*\vec{H}_0]\}, \vec{n})\, ds \tag{6}$$

($\vec{E}_0^*$, $\vec{H}_0^*$ is the complex conjugate of $\vec{E}_0$, $\vec{H}_0$), we find that the amplitude C_{-0} of the wave received by the antenna is equal, within the feed to

$$C_{-0} = \frac{\int_{s_1} (\{[\vec{E}^n\vec{H}_0] - [\vec{E}_0\vec{H}^n]\}, \vec{n})\, ds}{\int_{s_2} (\{[\vec{E}_0\vec{H}_0^*] + [\vec{E}_0^*\vec{H}_0]\}, \vec{n})\, ds}. \tag{7}$$

The $\vec{E}_{-0}$, $\vec{H}_{-0}$ wave received by the antenna is similar to the $\vec{E}_0$,$\vec{H}_0$ wave but moves in the opposite direction.

Let us also write the expression for the received power:

$$P_{\text{пр}} = \frac{\left| \int\limits_{s_1} (\{[\vec{E}^n \vec{H}_0] - [\vec{E}_0 \vec{H}^n]\}, \vec{n})\, ds \right|^2}{4 \int\limits_{s_2} (\{[\vec{E}_0 \vec{H}_0^*] + [\vec{E}_0^* \vec{H}_0]\}, \vec{n})\, ds} . \tag{8}$$

Formulas (6) and (7) are valid also in the case when within the feed there may be propagated M mutually orthogonal wave modes, creating different radiation fields $\vec{E}_m$, $\vec{H}_m$ (m<M). Formulas (6) and (7) in this case yield the field and power of an m-type wave in the channel. These fields and the residue, which is converted by the antenna into a scattered field, are mutually orthogonal; the total dissipation cross section of the antenna is the sum of the dissipation cross sections for each component of the principal field.

In conclusion it is my pleasant duty to express my sincere thanks to A. A. Pistol'kors and Ya. N. Fel'd for their valuable advice in the preparation of this work.

REFERENCES

1. E. L. Burshteyn, Concerning the power received by an antenna upon the incidence of a nonplane wave, Radiotekhnika i elektronika, 1958, 3, 2, 186.
2. Ya. N. Fel'd, L.S. Benenson, Antenny santimetrovykh i detsimetrovykh voln,[Antennas for centimeter and decimeter wavelengths] Izd. VVIA, im. N. Ye Zhukovskiy, 1955.
3. L.A. Vaynshteyn, Elektromagnitnyye volny [Electromagnetic waves]; Iazd, Sovetskoye radio, 1957.

Submitted to the editors 10 November 1960

ATTENUATION OF SLOW ELECTROMAGNETIC WAVES IN A PLASMA ROD IN A LONGITUDINAL MAGNETIC FIELD

V. P. Shestopalov and I. P. Yakimenko

The dispersion equation for a plasma rod in a longitudinal magnetic field was determined and examined without consideration of loss in reference [1]. This same equation was derived in reference [2] by the limit transition from a plasma-helix system to a plasma rod and has the following form:

$$\varepsilon_z f \frac{I_{11} I_{12}}{I_{01} I_{02}} + \frac{\sqrt{f_0}}{2 f_1 \sqrt{\varepsilon}} \left\{ [\varepsilon_z - 1 + (\varepsilon_z + 1) f_1] f_+ \frac{I_{11}}{I_{01}} - \right.$$
$$\left. - [\varepsilon_z - 1 - (\varepsilon_z + 1) f_1] f_- \frac{I_{12}}{I_{01}} \right\} \frac{K_{10}}{K_{00}} + \frac{f_- f_+}{|\varepsilon|} \frac{K_{10}^2}{K_{00}^2} = 0, \tag{1}$$

where

$$f_0 = 1 - \frac{1}{m^2}; \quad f_1 = \sqrt{1 + \frac{4\varepsilon_z}{m^2 \sigma^2 f_0^2}}; \quad f_\pm = \sqrt{\varepsilon_z \left(1 - \frac{\varepsilon}{m^2}\right) + \frac{\varepsilon - \varepsilon_z}{2} f_0 (1 \pm f_1)}; \tag{2}$$

$$I_{ik} = I_i(\gamma_k a); \quad K_{ik} = K_i(\gamma_k a) \ (i = 0,1; \ k = 0, 1, 2); \quad \gamma_0 = \beta \sqrt{f_0}; \quad \gamma_{1,2} = \frac{\beta f_\pm}{\sqrt{\varepsilon}};$$

$\beta = \omega/v_f$ is the longitudinal propagation constant; $m = c/v_f$ is the delay; ϵ and ϵ_z are the tensor components of the dielectric constant; $\sigma = \omega_H/\omega$; ω is the frequency of the signal; ω_H is the gyrofrequency; v_f is the phase velocity of the delayed wave; c is the velocity of light in vacuum; a is the radius of the plasma rod; $I_0(x)$ and $K_0(x)$ are modified Bessel functions of the first and second kinds.

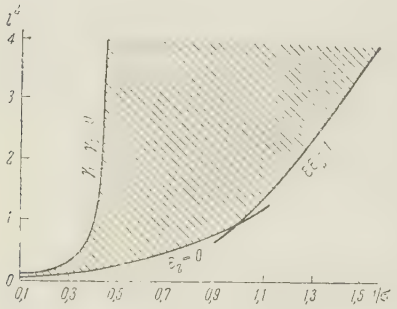


Figure 1

Equation (1) permits determination of the dispersion of slow electromagnetic waves. It is applicable within the dark area in Figure 1. This region is limited by the curves $\epsilon_z = 0$ $\epsilon\epsilon_z = 1$ (the boundary frequency, that is the frequency at which $m^2 \to \infty$.) and $(l\sigma)^2 = 2(\sigma - 1)/(\sigma - 2)$ ($l = \omega_0/\omega$, is the Langmuir plasma frequency). Below the latter curve γ_1 and γ_2 are simultaneously real numbers.

If attenuation of the waves is not extremely grea (that is, if the real parts of the tensor components of the dielectric constant ϵ', ϵ_z' and the longitudinal propagation constant β' are considerably greater tha the corresponding imaginary parts, ϵ'', ϵ_z'', β''), then expansion of the corresponding terms in Equatio (1) in terms of a small parameter with retention onl of linear quantities permits obtaining both the dispersion equation of the system (for the case of absence of loss) and its attenuation.

With large delays ($m > 10$) the dispersion equation takes the following relatively simple form:

$$\sqrt{\varepsilon'\varepsilon_z'} = \frac{I_0\left(\beta'\sqrt{\frac{\varepsilon_z'}{\varepsilon'}}\,a\right) K_1(\beta' a)}{I_1\left(\beta'\sqrt{\frac{\varepsilon_z'}{\varepsilon'}}\,a\right) K_0(\beta'\alpha)}, \tag{3}$$

and the losses in the system are defined by the equation

$$\beta'' a = \frac{\frac{\varepsilon'\varepsilon_z'' + \varepsilon_z'\varepsilon''}{2\sqrt{\varepsilon'\varepsilon_z'}}\frac{I_1}{I_0} + \frac{\beta' a}{2|\varepsilon'|}(\varepsilon_z''\varepsilon' - \varepsilon''\varepsilon_z')\left(\frac{I_1}{I_0}\right)'}{\left(\frac{K_1}{K_0}\right)' - |\varepsilon_z'|\left(\frac{I_1}{I_0}\right)'} \tag{4}$$

where

$$\varepsilon' = 1 + \frac{(l\sigma)^2}{\sigma^2 - 1}; \qquad \varepsilon'' = \frac{\left(1 + \frac{1}{\sigma^2}\right) l^2\sigma\delta}{\left(1 - \frac{1}{\sigma^2}\right)^2}; \tag{5}$$

$$\varepsilon_z' = 1 - (l\sigma)^2; \qquad \varepsilon_z'' = l^2\sigma^3\delta; \quad \delta = \frac{\nu}{\omega_H};$$

ν is the effective collision frequency within the plasma; $(I_1 / I_0)'$ and $(K_1 / K_0)'$ are the argument derivatives of the Bessel function ratios. The arguments of the functions I_i are equal to $\beta'\sqrt{\frac{\varepsilon_z'}{\varepsilon'}}\,a$; the arguments of the function K_1 is $\beta' a$. The expressions (5) are correct if the operating frequency is remote from the gyrofrequency ($\sigma \neq 1$).

Dispersion dependence of the plasma rod and attenuation are shown in Figures 2 and 3 for $l^2 = 3\omega_H a/c = 0.1$ and $\delta = 10^{-3}$. The curves show that in a plasma rod located in a longitudinal magnetic field we may derive the delay of axially symmetrical waves on the order of 10-20 with attenuations below 0.1 db/cm. With further approximation to the boundary frequency, simultaneously with an increase in delay there is a marked increase in losses.

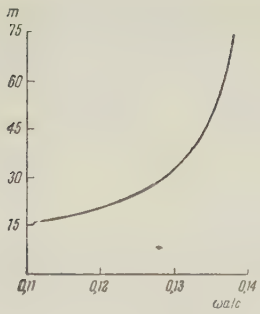


Figure 2

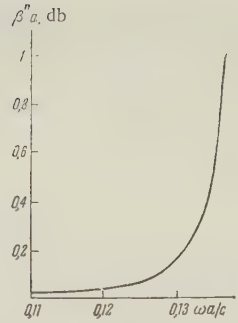


Figure 3

REFERENCES

1. Ya.B. Faynberg, M.F. Gorbatenko, ZhTF, 1959, 29, 5, 549.
2. B.M. Bulgakov, V.P. Shestopalov, L.A. Sheshkin, I.P. Yakimenko, ZhTF, 1960, 30 7, 805.

Khar'kov State University
imeni A.M. Gor'kiy

Submitted to the editors 9 November 1960

ON ELECTRODYNAMIC BOUNDARY CONDITIONS AT A PLANE SURFACE WITH ARBITRARY VALUE OF DIELECTRIC CONSTANT

F.G. Bass

In reference [1] the boundary condition was derived for the normal component of the electric field at a plane surface with arbitrary value of dielectric constant. In a number of cases this is not a fully defined solution of the problem. For example, for fields in which the normal component is an identical equality with zero this condition becomes an identity and the boundary problem must be solved by the use of the boundary conditions for the tangential components of the electromagnetic field.

The boundary conditions for the tangential components of the electric field are obtained by means of Green's vector formula and the precise boundary conditions by a method similar to that which was used in reference [1]. These boundary conditions have the form

$$[\vec{n},\ \vec{E}(\vec{r})] = -\frac{ik}{2\pi}\int \frac{e^{ik\sqrt{\varepsilon}|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}\left\{[\vec{n}\,[\vec{n},\ \vec{H}(\vec{r}')]] + \right.$$
$$\left. + \frac{1}{k^2\varepsilon}[\vec{n}\nabla](\nabla[\vec{n},\ \vec{H}(\vec{r}')])\right\} d\vec{r}' \quad \text{for } z = 0. \tag{1}$$

Here $\vec{E}$ and $\vec{H}$ are the electric and magnetic fields; $k = \omega/c$; ϵ is the permittivity; $\vec{n}$ is the exterior normal to the plane $z = 0$.

If $|\epsilon| > 1$, formula (1) may be rewritten as

$$[\vec{n},\ \vec{E}] = \frac{1}{\sqrt{\varepsilon}}\left(1 + \frac{1}{k^2\varepsilon}\Delta\right)^{-1/2}\{[\vec{n}\,[\vec{n},\ \vec{H}]] + \frac{1}{k^2\varepsilon}[\vec{n}\nabla](\nabla[\vec{n},\ \vec{H}])\} \quad \text{for } z = 0, \tag{2}$$

where Δ is the bidimensional Laplace operator in coordinates x and y.

Using the Maxwell equations, we may eliminate the magnetic field from (2). After simple transformations we obtain these boundary conditions for the electric field:

$$\frac{\partial E_{x,y}}{\partial z} = ik\sqrt{\varepsilon}\left(1+\frac{1}{k^2\varepsilon}\Delta\right)^{1/2}E_{x,y} = \frac{\varepsilon-1}{\varepsilon}\frac{\partial E_z}{\partial x,y}, \tag{3}$$

for $z = 0$

$$\frac{\partial E_z}{\partial z} = -\frac{ik}{\sqrt{\varepsilon}}\left(1+\frac{1}{k^2\varepsilon}\Delta\right)^{1/2}E_z.$$

We will note that the second of the boundary conditions in (3) is the direct result of boundary conditions in the form of (2) or, which is the same thing, the first of the boundary conditions in (3).

Expanding boundary condition (2) in powers of $\Delta/k^2\epsilon$ and retaining the first term of expansion, we obtain the Leontovich boundary conditions.

Using the boundary conditions of (3), let us solve the following problem. Consider the class of solutions of homogeneous wave equations* for which boundary conditions (1) - (3) are local. It is easily seen that plane waves are included in this class. The boundary problem for plane waves is solved by means of the reflection coefficient. It is found that it may be introduced in the more general case.

From the boundary conditions in the form of (2)-(3) it is seen that they are local if the components of the electromagnetic fields at the plane $z = 0$ are eigenfunctions of the bidimensional Laplace operator Δ. We shall limit the discussion to the less general case in which the components of the electromagnetic field are eigenfunctions of the bidimensional Laplace operator in the entire upper half space. As is easily seen, solution of the wave equation on this assumption may be sought in the following form:

$$E_i = g_i^{(+)}(x, y)\,e^{i\varkappa z} + g_i^{(-)}(x, y)\,e^{-i\varkappa z}, \tag{4}$$

where $g_i^{(\pm)}(x,y)$ satisfies the equation

$$\Delta g_i^{(\pm)}(x, y) + (k^2 - \varkappa^2)\,g_i^{(\pm)} = 0 \quad (i = x, y, z). \tag{5}$$

As is known (reference [2]), the general solution of equation (5) is written as

$$g(x, y) = \alpha J_0(\lambda\rho) + \int_0^{z_+} \Phi_1(t)\,J_0\left(\lambda\sqrt{z_-(z_+ - t)}\right)dt + \int_0^{z_-} \Phi_2(t)\,J_0\left(\lambda\sqrt{z_+(z_- - t)}\right)dt, \tag{6}$$

where α is an arbitrary constant $z_\pm = x \pm iy$; $\lambda = \sqrt{k^2 - \varkappa^2}$; $\rho = |z_\pm|$; Φ_1 and Φ_2 are arbitrary holomorphic functions; J_0 is a zero order Bessel function. If E_z is not identically equal to zero the solution of the Maxwell equations has the form

$$E_{x,y} = \frac{i\varkappa}{\varkappa^2 - k^2}\left\{e^{i\varkappa z} - R(\varkappa)\,e^{i\varkappa z}\right\}\frac{\partial g_z}{\partial x, y},$$

$$E_z = \left(e^{-i\varkappa z} + R(\varkappa)\,e^{i\varkappa z}\right)g_z, \qquad R(\varkappa) = \frac{1 - \frac{k}{\varkappa\sqrt{\varepsilon}}\left(1+\frac{\varkappa^2-k^2}{k^2\varepsilon}\right)^{1/2}}{1 + \frac{k}{\varkappa\sqrt{\varepsilon}}\left(1+\frac{\varkappa^2-k^2}{k^2\varepsilon}\right)^{1/2}}. \tag{7}$$

If E_z is identically equal to zero,

$$E_{x,y} = \left(e^{-i\varkappa z} + P(\varkappa)\,e^{i\varkappa z}\right)g_{x,y}, \qquad P(\varkappa) = \frac{1 - \frac{k}{\varkappa}\sqrt{\varepsilon}\left(1+\frac{\varkappa^2-k^2}{k^2\varepsilon}\right)^{1/2}}{1 + \frac{k}{\varkappa}\sqrt{\varepsilon}\left(1+\frac{\varkappa^2-k^2}{k^2\varepsilon}\right)^{1/2}} \tag{8}$$

For plane waves formulas (6) and (7) change to the ordinary formulas for vertical and horizontal polarization. Quantities $R(\varkappa)$ and $P(\varkappa)$ have the meaning of reflection coefficients. As is seen from formula (6), the class of solutions for which the boundary problem is

*In the discussed problem the Maxwell equations reduce to wave equations for the components of an electromagnetic field.

solved by introduction of the reflection coefficients is not limited to plane waves, for even in the discussed special case it is identified by two arbitrary functions. Of course, the problem of excitation of a field described by specific α, Φ_1 and Φ_2, must be solved separately in each case.

Let us note that, knowing the solutions of the homogeneous wave equation with boundary conditions (1) - (3), we may solve a wave equation with the right-hand member by ordinary methods of mathematical physics.

REFERENCES

1. F.G. Bass, Radiotekhnika i elektronika, 1960, 5, 3, 389.
2. I.N. Bekua, Novyye metody resheniya ellipticheskikh uravneniy [New methods for solution of elliptic equations]; GTI, 1948.

Institute of Radiophysics and
Electronics
AN USSR

Submitted to the editors 30 April 1960

Note Added in Proof

It is easily shown that the obtained boundary conditions also hold for a nonuniform surface and for ϵ, depending on the coordinates, if the following inequalities obtain:

$$\left|\frac{\nabla\varepsilon}{k\varepsilon^{3/2}}\right| \ll 1, \quad R \gg \left|\frac{1}{k\sqrt{\varepsilon}}\right|,$$

where R is the radius of curvature of the surface.

EVALUATING THE INFLUENCE OF THERMAL VELOCITIES IN ELECTRON BEAMS

L.E. Bakhrakh

From the standpoint of a simplified method proposed below, we shall discuss ion focusing of an electron stream and the Brillouin stream taking into account thermal velocities.

1. ION-FOCUSED BEAM

Ion focusing of electron beams has been discussed in a number of works (references [1-3]. Essentially, this type of focusing consists in the fact that ions originating in a gas during the transit of an electron beam accumulate in the path of the latter, forming a positive space-charge strand which attracts electrons toward the axis and thus focuses the electron stream.

For an ion-focused cylindrical electron beam of radius r moving along an axis z the equatic of radial motion of the boundary electron may be presented in the form

$$m\frac{d^2r}{dt^2} = -f_+ + f_- + p_1,$$

where f_+ is the focusing force of the positive ions, f_- is the defocusing force of Coulomb repulsion of the beam electrons, p_1 is the force resulting from transverse thermal velocities of the electrons in the stream, m is the electron mass.

On the basis of the above assumptions we may easily determine f_+ and f_- by using the Gauss theorem:

$$f_+ = -e\frac{\rho_+ r}{2\varepsilon_0}, \quad f_- = e\frac{\rho_- r}{2\varepsilon_0} = e\frac{I}{2\pi r v \varepsilon_0}, \tag{1}$$

where $\rho+$ and $\rho-$ are the mean densities of the positive ions and electrons in the beam, I is the beam current intensity, $v = \sqrt{\frac{2e}{m}U}$ is the axial velocity of the electrons, U is the accelerating voltage, e is the electron charge, ϵ_0 is the dielectric constant of the vacuum.

The scattering force p_1 caused by thermal movement of electrons may be treated as follows. Let us assume that the pressure of the electron gas P = nkT, where n is the concentration of electrons in the stream, k is the Boltzman constant, T is the absolute temperature. Then the force acting on the surface of a beam of radius r and length l wil be equal to $P2\pi r\, l$. Expressing this force in terms of one electron, we obtain

$$p_1 = \frac{2kT}{r}. \tag{2}$$

We will note that the introduction of force p_1 as defined in Equation (2) may be justified to some degree if we regard the electron beam as a certain system of N physical points the equilibrium conditions of which, under the influence of active forces F and inertial forces $\bar{f}_i$, is defined by the d'Alembert-Lagranve equation

$$\sum_{i=1}^{N} (\bar{F}_i - \bar{f}_i)\,\delta\bar{r}_i = 0.$$

In our case of a cylindrical beam of surface S under the influence of forces caused by the transverse thermal velocities of electrons we may assume that

$$\sum_{i=1}^{N} \bar{F}_i \delta\bar{r}_i = PS\delta r.$$

If we also assume that $\bar{f}_i$ and $\delta\bar{r}_i$ are identical for all points of the system, then

$$\sum_{i=1}^{N} \bar{f}_i \delta\bar{r}_i = nVf\delta r = PS\delta r,$$

where V is the volume of the investigated portion of the stream.

Hence,we have the relation, similar to (2):

$$f = \frac{PS}{nV} = \frac{2kT}{r}.$$

Let us adopt an adiabatic expression as the equation of state of the gas and use the condition of adiabaticity (reference [2]), which in this case has the form

$$\frac{\varphi_T}{n} = \text{const},$$

where $\varphi_T = kT/e$. Since n is proportional to r^{-2}, we may write

$$\varphi_T r^2 = \varphi_{TK} r_K^2 = \text{const}, \tag{3}$$

where $\varphi_{TK} = kT_K/e$; T_K is the cathode temperature and r is the radius defining the position of the electron at the cathode.

Keeping in mind relations (1), (2) and (3), the equation of motion of the boundary electron will be

$$m\frac{d^2r}{dt^2} = -\frac{e\rho_+ r}{2\varepsilon_0} + \frac{cI}{2\pi r v \varepsilon_0} + e\frac{2\varphi_{TK} r_K^2}{r^3}.$$

Assuming v = const, this equation may be transformed to

$$\frac{d^2r}{dz^2} = -\frac{\rho_+ r}{4U\varepsilon_0} + \frac{I}{4\pi r U \varepsilon_0} + \frac{\varphi_{TK} r_K^2}{Ur^3}. \tag{4}$$

Equation (4) wholly coincides with the differential equation for effective radius of a stream with ion focusing obtained by a more rigorous method in reference [2].

It must also be pointed out that application of the described method of investigation to the case of ion focusing of a flat beam leads to results similar to those obtained by the more rigorous method of calculation.

2. BRILLOUIN STREAM

Let us discuss the problem of the influence of thermal velocities in a Brillouin stream of electrons moving along axis z. As is known (reference [4]), the equation of radial motion of a boundary electron of the beam in this case has the form

$$\frac{d^2r}{dz^2} = \eta \frac{I}{2\pi\varepsilon_0 r v^3} - r\left(\frac{\eta}{2}\frac{B}{v}\right)^2, \tag{5}$$

where I is the beam current intensity, v is its velocity, B is the magnetic induction, η is the charge-to-mass ratio of the electron. The equilibrium Brillouin radius may be found from the condition $d^2r / dz^2 = 0$ and is

$$r_b^2 = \frac{\sqrt{2}\, I}{2\pi\varepsilon_0 \eta^{3/2} B^2 U^{1/2}}.$$

In order to evaluate the influence of thermal velocities on focusing of the beam let us add to the right-hand side of Equation (5) the defocusing force $f = 2kT/r$ caused by the presence of the thermal velocities. As before, adopting the adiabatic equation as the equation of the state of the gas and using the condition of adiabaticity $\varphi_T/n = \text{const}$, we obtain the equation of motion of the boundary electron of the beam in the form:

$$\frac{d^2r}{dz^2} = \eta \frac{I}{2\pi\varepsilon_0 r v^3} - r\left(\frac{\eta}{2}\frac{B}{v}\right)^2 + \frac{\varphi_{T\,K} r_K^2}{U r^3} \tag{6}$$

If in Equation (6) we assume that $d^2r / dz^2 = 0$, we may determine the so-called equilibrium radius r_m of the stream taking into account the influence of thermal velocities. For this radius we obtain

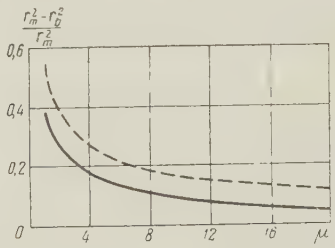
$$r_m^4 - r_b^2 r_m^2 - \frac{1}{\mu} r_b^4 = 0,$$

where

$$\mu = \frac{Ie}{kT_K v}\frac{r_b^2}{r_K^2} = 1.76\cdot 10^8 \frac{I}{T_K \sqrt{U}}\frac{r_b^2}{r_K^2}.$$

From this last relation it follows that

$$r_m^2 = \frac{r_b^2}{2}\left(1 + \sqrt{1 + \frac{4}{\mu}}\right). \tag{7}$$



On the assumption of uniform electron density in the beam, Equation (7) permits an estimation of that part of the total charge of the beam which passes beyond the Brillouin radius due to the scattering action of thermal velocities. This part of the beam charge is evidently

$$\frac{r_m^2 - r_b^2}{r_m^2} = 1 - \frac{1}{\frac{1}{2}\left(1 + \sqrt{1 + \frac{4}{\mu}}\right)}.$$

On the basis of Equation (8) the solid curve in the figure represents the dependence of the part of the beam current lying beyond the Brillouin radius on the value of parameter μ. The dashed curve in the figure represents the data of the more rigorous calculation of Pierce and Walker. Comparison the these curves shows satisfactory agreement.

Investigation of an ion-focused stream and a Brillouin stream permits the assumption that the described method of evaluating the influence of thermal velocities of electrons within the beams may be applied in a number of other cases.

I wish to thank P.V. Golubkov and A. M. Aleskovskiy for their valuable comments.

REFERENCES

1. O. Scherzer, Zh. Phys., 1933, 82, 6, 693.
2. B.I. Davydov, S.I. Braginskiy. Toward a theory of gas concentration of electron beams.

Sb., posvayashchennyy 70-letiyu akademika A.F. Ioffe, [Symposium in honor of Academician A.F. Ioffe on his seventieth birthday]; Izd. AN SSSR, 1950, 72-91.
3. M.M. Bredov, Automatic compensation of space charge in electron beams, Sb., posvyashchennyy 70-letiyu akademika A.F. Ioffe, Izd. AN SSSR, 1950, 155-172.
4. J.R. Pierce, Theory and calculation of electron beams, transl, from English, Izd. Sovetskoye radio, 1956, 154-160.
5. J. Pierce, L. Walker, "Brillouin flow" with thermal velocities, J. Appl. Phys., 1953, 24, 10, 1328.

Submitted to the editors 25 April 1960

CALCULATION OF HOLLOW ELECTRON BEAMS

N.D. Porev

In the transition from solid cylindrical electron beams to hollow beams three questions usually arise.

1. To what extent do the repulsive forces change and, consequently, the values of magnetic flux density required for shaping?

2. How much more efficient does interaction of the beam with the microwave field become?

3. How does the current density of emission from the cathode vary?

Let us assume that the beam current I and outer radius r_B are constant and investigate the change in properties of the beam with a change in radius of the internal cavity r_a and, consequently, in the "hollowness" $H = r_a/r_b$.

1. The repulsive forces acting on the boundary electron have been investigated more fully in reference [1]. However, the use of series and consideration of second-order effects render the results of this work relatively inaccessible. Moreover, with the usual omissions adequate results may be obtained by assuming that the space-charge density does not vary over the cross-section of the beam.

In this case the repulsive force acting on a boundary electron is represented by the expression

$$F_r = \frac{e\rho}{2\varepsilon r_B}(r_B^2 - r_a^2) = \frac{eI}{2\pi\varepsilon z r_B}, \tag{1}$$

that is, with a given current the repulsive force acting on an outer boundary electron and, consequently, the magnetic field required for shaping do not depend on the hollowness of the beam

2. Interaction of the beam with the microwave field is adequately characterized by the gain parameter (reference [2])

$$C = \sqrt[3]{\frac{KI}{4U_0}}. \tag{2}$$

In the first approximation the change in the axial field component as a function of the radius, in the presence of a beam, may be calculated, as for a "cold" delay system, by the corresponding expression

$$E_z(r) = AJ_0(\gamma r). \tag{3}$$

Considering the relations (reference [2]

$$K = \frac{E_z^2}{2\beta^2 P}, \tag{4}$$

$$C^3 = \frac{IE_z^2}{8U_0\beta^2 P} = \frac{A^2 I J_0^2(\gamma r)}{8U_0\beta^2 P} \tag{5}$$

and dividing the beam into annular elements with constant $E_Z(r)$ and current densities $j = \rho\dot{z}$, we obtain

$$C^3 = \frac{A^2}{8U_0\beta^2 P} \int_{r_a}^{r_B} J_0^2(\gamma r)\, \rho\dot{z} 2\pi r\, dr. \tag{6}$$

The condition of equality of currents requires

$$I = (r_B^2 - r_a^2)\pi\rho\dot{z} = \text{const}. \tag{7}$$

For a solid beam $r_a = 0$. Then

$$\frac{B_{HB}^3}{C_{SB}^3} = \frac{\rho_T \int_{r_a}^{r_B} J_0^2(\gamma r)\, r\, dr}{\rho_{SB} \int_0^{r_B} J_0^2(\gamma r)\, r\, dr} = \frac{r_B^2}{r_B^2 - r_a^2} \frac{\int_{r_a}^{r_B} J_0^2(\gamma r)\, r\, dr}{\int_0^{r_B} J_0^2(\gamma r)\, r\, dr}. \tag{8}$$

Let us evaluate the results in regions where approximations of the modified cylindrical functions are applicable. With small $\gamma r J_0(\gamma r) \simeq 1$ and

$$\frac{C_T^3}{C_{CH}^3} = \frac{r_B^2}{r_B^2 - r_a^2} \frac{\int_{r_a}^{r_B} r\, dr}{\int_0^{r_B} r\, dr} = 1, \tag{9}$$

which can be well interpreted physically, since the microwave field in this approximation does not vary within the confines of the beam.

For large $\gamma r J_0(\gamma r) = e^{\gamma r}/\sqrt{2\pi\gamma r}$ and

$$\frac{C_{HB}^3}{C_{SB}^3} = \frac{r_B^2}{r_B^2 - r_a^2} \frac{\int_{r_a}^{r_B} e^{2\gamma r}\, dr}{\int_0^{r_B} e^{2\gamma r}\, dr} = \frac{1}{1 - T^2} \frac{e^{2\gamma r_L} - e^{2\gamma r_a}}{e^{2\gamma r_B} - 1}. \tag{10}$$

It would be more rigorous to use two approximations within the integration limits by dividing the integral into two parts, but this has no substantial effect on the result.

3. The area of the cathode within a hollow beam is decreased by

$$\frac{\pi r_B^2}{\pi (r_B^2 - r_a^2)} = \frac{1}{1 - T^2} \tag{11}$$

times, while the emission current density is increased by the same factor.

Numerical calculations for a number of cases of practical interest ($2\gamma r_b \simeq 6.5$) yield the following values:

HB	0.4	0.6	0.8	0.9
$\frac{C_{HB}}{C_{SB}}$	1.08	1.14	1.27	1.37
$\frac{j_{HB}}{j_{SB}}$	1.19	1.56	2.78	5.27

A substantial increase in gain parameter occurs only with a highly hollow beam, but the required emission current density also increases sharply. Hence, from this standpoint it is convenient to use only highly hollow beams, although this is possible only with sufficiently high cathode emissivity.

REFERENCES

1. L.A. Harris, Proc. IRE, 1952, 40, 6, 700.
2. J. R. Pierce, Traveling-wave tube, transl. from English under editorial supervision of V.T. Ovcharov, Izd. Sovetskoye radio, 1952.

Submitted to the editors 7 July 1960

POLARIZATION ANALYZER

V.A. Libin

INTRODUCTION

In investigations of the polarization characteristics of the electromagnetic field of antenna systems, individual radiators, polarization gratings, elements of waveguide systems, etc. and also in studying the effects of a change in the polarization parameters of a field during propagation through an anisotropic medium and in a number of other cases it is necessary to determine the parameters of the polarization ellipse (the power coefficient of ellipticity M and the orientation angle β of the major axis) from the known characteristics of two orthogonal fields --their relative phases and amplitudes. In the solution of such problems it is usually necessary to perform calculations or to use the chart described in reference [1]. Although the second method is preferred this chart has certain shortcomings (e.g., the need for use of an auxiliary straightedge). In addition, it is convenient to have a diagram for polarization analysis of the same type as the polarization-matching chart described in reference [2], that is, in the form of a polar diagram. This permits combining both diagrams in one and this may be used for the solution of most practical polarization problems.

The chart described below (known as a polarization analyzer) is made in the form of a sector of a circle and should be used as a supplement to the polarization-matching diagram.

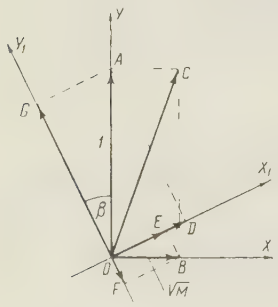


Figure 1

1. DERIVATION OF RELATIONS BETWEEN POLARIZATION PARAMETERS

Let there be an electromagnetic field the ellipticity of which (relative to power) is M. The vectors of electric field strength, parallel to the major and minor axes of the polarization ellipse, are conditionally shown in phase in Figure 1. The amplitude values of vectors $\vec{OA}$ and $\vec{OB}$, referred to $|\vec{OA}|$, are 1 and $\sqrt{M}$, respectively. The instantaneous values of these vectors (here and hereafter normalized relative to $|\vec{OA}|$) are:

$$|OA| = \sin \omega t, \quad |OB| = \sqrt{M} \cos \omega t$$

(in the reverse direction of field rotation $|OA| = \cos \omega t$, , $|OB| = \sqrt{M} \sin \omega t$.)

In order to determine the relation between the relative amplitudes and phases of the two orthogonal components of this field (with these components displaced by an angle β relative to the axes of the polarization ellipse) and between the ellipticity M and angle β it is necessary to find the sums of the projections of these components onto the axes of the polarization ellipse or, conversely, to project the vectors of the fields parallel to the major and minor axes of the ellipse onto the directions subtending the angle β with these axes.

Having found the sums of the projections of these vectors, we may determine their relative amplitudes and phases.

The projections onto directions Y_1 and X_1 are

$$OG = |OA| \cos \beta = \cos \beta \sin \omega t, \quad OF = -|OB| \sin \beta = -\sqrt{M} \sin \omega t,$$
$$OD = |OA| \sin \beta = \sin \beta \sin \omega t, \quad OE = |OB| \cos \beta = \sqrt{M} \cos \beta \cos \omega t.$$

The sum of the projections onto axis X_1 are

$$OD + OE = \sin \beta \sin \omega t + \sqrt{M} \cos \beta \cos \omega t = A_1 \sin (\omega t + \varphi_1),$$

where

$$A_1 = \sqrt{\sin^2 \beta + M \cos^2 \beta}; \quad \varphi_1 = \operatorname{arc\,tg} [\sqrt{M} \operatorname{ctg} \beta].$$

The sum of the projections onto axis Y_1 are

$$OG + OF = \cos \beta \sin \omega t - \sqrt{M} \sin \beta \cos \omega t = A_2 \sin (\omega t + \varphi_2),$$

where

$$A_2 = \sqrt{\cos^2 \beta + M \sin^2 \beta}; \quad \varphi_2 = \operatorname{arc\,tg} [-\sqrt{M} \operatorname{tg} \beta].$$

The relative amplitude of the sums of the projections are

$$\frac{A_1}{A_2} = \sqrt{\frac{\sin^2 \beta + M \cos^2 \beta}{\cos^2 \beta + M \sin^2 \beta}}.$$

The ratio of powers carried by the corresponding fields is

$$N = \frac{P_1}{P_2} = \left[\frac{A_1}{A_2}\right]^2 = \frac{\sin^2 \beta + M \cos^2 \beta}{\cos^2 \beta + M \sin^2 \beta}. \tag{1}$$

The phase shift between the orthogonal components is

$$\Delta\varphi = \varphi_1 - \varphi_2 = \operatorname{arc\,tg} [\sqrt{M} \operatorname{ctg} \beta] - \operatorname{arctg} [-\sqrt{M} \operatorname{tg} \beta] = \operatorname{arc\,tg} \left[\frac{2\sqrt{M}}{(1-M) \sin 2\beta}\right]. \tag{2}$$

2. USE OF THE POLARIZATION ANALYZER

For each of Equations (1) and (2) there is a corresponding family of curves. Simultaneous solution of (1) and (2) with two parameters given, for example, N and $\Delta\varphi$, yields the desired values of M and β.

In Figure 2 the curves for formulas (1) and (2) are plotted in polar coordinates. In both families of curves the ellipticity M is the chosen parameter. The value of N (or $\Delta\varphi$) is determined from the length of the radius-vector extended at a given angle β to the curve corresponding to the given value of M.

From (1) it is easily seen that $N(\beta) = \frac{1}{N}\left(\frac{\pi}{2} - \beta\right)$, whence it follows that the diagram may be plotted in sector $\pi/4$, wherein for angles $\beta < \pi/4$ it is necessary to choose reciprocal values of N. A similar conclusion may be made concerning (2), the only difference being that for $\beta < \pi/4$ the phase angle $\Delta\varphi$ will be negative.

For rapid determination of the required values the polarization analyzer has a pivotable slide with scales for $\Delta\varphi$ and N.

The analyzer is used as follows.

1. If β and M are given: (a) rotate the slide through angle β; (b) find the point of intersection of the slide with that curve from the $\Delta\varphi$ family which corresponds to the given value of M; (c) read the value of $\Delta\varphi$ on the slide scale; (d) N is found similarly.

2. If N and $\Delta\varphi$ are given: (a) find on the slide the points corresponding to the given values of N and $\Delta\varphi$; (b) turn the scale until these points coincide with the curves of N and

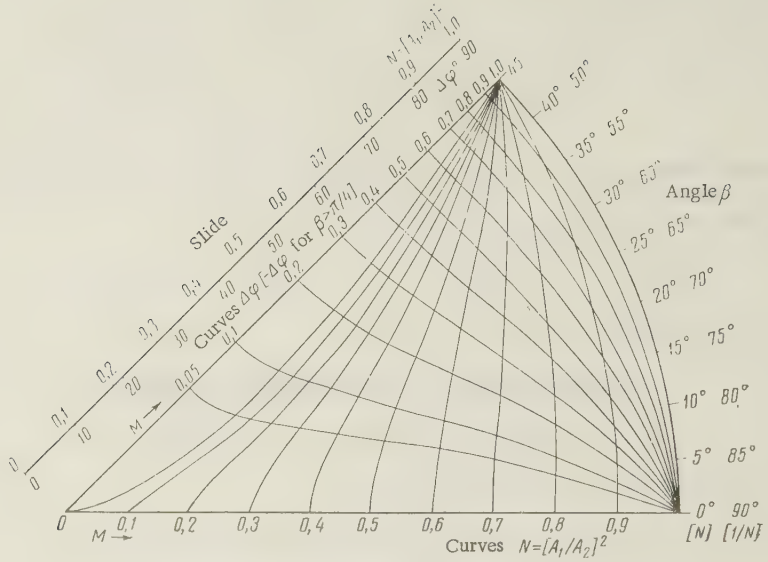


Figure 2

$\Delta\omega$, respectively, plotted for the same value of parameter M; (c) the resulting values of β and M are the sought values. Given any two of these quantities, it is not difficult to determine the other two.

It was stated above that it would be desirable to combine the polarization matching diagra with the described polarization analyzer. Figure 3 gives the scheme for construction of such a diagram. In the figure: 1, polarization matching diagram; 2, families of auxiliary curves for matched polarizations; 3, the same for mismatched polarizations; 4, polarization analyzer; 5, pivoted slide scale.

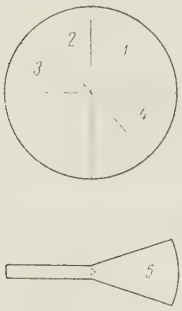


Figure 3

REFERENCES

1. M. Morgan, W. Evans, Proc. IRE, 1951, 39, 552.
2. V. A. Libin, Radiotekhnika i elektronika, 1960, 5, 11, 1786.

Submitted to the editors
7 October 1960

POSSIBLE APPLICATION OF FERRITES FOR ABSOLUTE MEASUREMENT OF MICROWAVE POWER

G. B. Bogdanov

It is the purpose of the present work to determine the properties of manganese ferrite as a ferrimagnetic thermistor which might be used for the absolute measurement of microwave power in bridge circuits.

The possibility of such application is based on dependence of the electrical resistance of ferrites on temperature and on the intense (above 100° C) heating of a ferrite specimen as the result of absorption of high levels of microwave power at ferrimagnetic resonance. The dependence of ferrite resistance on temperature is described by a function of the form (reference [1])

$$R_f = Ae^{\Delta E/kT}, \tag{1}$$

where ΔE is the activation energy, k is the Boltzman constant, T is absolute temperature, A is a constant coefficient expressed in ohms.

The activation energy is not a constant quantity and takes different values for various temperatures (reference [1]). Without dwelling on the microscopic processes occurring in the ferrite upon heating, the activation energy ΔE may be considered as a constant averaged quantity in the temperature range from room temperature (23° C) to the Curie point. This assumption is acceptable for engineering calculations for ordinary (not possessing ferrimagnetic properties) thermistors (reference [2]). It is shown below that this assumption remains valid for ferrimagnetic semiconductor thermistors. With this assumption formula (1) may be presented in the form

where
$$R_f = Ae^{B/T}, \qquad B = \frac{\Delta E_{av}}{k} = \text{const}, \tag{2}$$

that is, coefficient B characterizes the ratio of the averaged quantity of activation energy ΔE_{av} to the Boltzmann constant and is expressed in degrees Kelvin.

Since in the absorption of high levels of microwave power at ferrimagnetic resonance the ferrite is intensely heated, the ferrite resistance will decrease with an increase in the absorbed power. The relation between electrical resistance of the ferrite and the absorbed microwave power will then by defined by relation

$$R_f = Ae^{B/T+cP_f}, \tag{3}$$

where c is a constant coefficient expressed in degrees Kelvin per watt, P_f is the power absorbed by the ferrite at ferrimagnetic resonance. The absorbed power P_f is determined from the relation

$$P_f = F\frac{\omega V_f M_0 h^2}{2\Delta H}, \tag{4}$$

where F is a coefficient depending on the shape of the ferrite specimen, ω is the frequency of the external magnetic microwave field, h is the amplitude of the external microwave field, M_0 is the saturation magnetization, V_f is the volume of the ferrite specimen, $2\Delta H$ is the width of the absorption line. It must be pointed out that in operation at high microwave powers the dependence of M_0 and $2\Delta H$ on temperature has considerable effect on the value of P_f. For example, the power absorbed at ferrimagnetic resonance at 23° C by manganese ferrite in the shape of a disc decreases by 20%-30% upon heating to 150° C. Hence, for precise calculation of P_f the temperature dependence of M_0 and $2\Delta H$ cannot be disregarded.

Manganese ferrite was chosen for experiment because it possesses higher conductivity

than the ferrites presently used in microwave engineering and hence it is more suitable for the purpose.

All experiments (except that for determining the dependence of ferrite resistance R_f on absorbed power P_f) were conducted with a cylindrical specimen whose length is 2.7 mm and diameter is 2.6 mm. The resistance of this specimen at room temperature was $1.1 \cdot 10^6$ ohms. Fine copper conductors leading to the device for measurement of ferrite resistance were soldered to thermal-bonded silver contacts at the endfaces of the cylinder.

Figure 1. Resistance of manganese ferrite as a function of temperature for different values of magnetizing field H_0.

In order to determine the dependence of ferrite resistance on absorbed power a ferrite disc with diameter of 6 mm. and thickness of 0.8 mm was used. At room temperature the resistance of this specimen was $1.55 \cdot 10^5$ ohms. The contacts were applied to the plane surfaces of the disc.

The cylindrical specimen was used for investigation of ferrite conductivity as a function of temperature T and a constant magnetizing field H_0 (galvanomagnetic effect) and also to obtain the static volt-ampere characteristic of the ferrite.

The results of experimental investigation of $R_f = f(T)$ with magnetizing field $H_0 = 0$ and $H_0 = 3500$ oersteds are shown in Figure 1.

The curves for $R_f = f(t)$ in Figure 1 are typical of semiconductor thermistors and are distinguished by the fact that at room temperature the ferrite conductivity is several orders of magnitude greater than in ordinary semiconductor thermistors. With the assumption that $\Delta E_{av}/k = B$ it may be shown that the curves for $R_f = f(T)$ in Figure 1 are satisfactorily described by formula (2) in the temperature range of 45 - 200° C if $A \simeq 1.52$ and $B \simeq 3950$. These coefficients are determined from the curves of $R_f = f(T)$ and depend on the material and dimensions of the ferrite specimen. Knowing the numberical value of coefficient B, we may determine the value of ΔE in the above-mentioned temperature range. ΔE_{av} will be 34 ev. For the sake of comparison we will note that ΔE_{av} as obtained by another means for manganese ferrite monocrystals (reference[1]) is approximately 27 ev in the same temperature range. The difference in values may be explained by differences in composition of the investigated ferrites and also by various errors in measurement.

This comparison of values of ΔE_{av} permits us to consider the assumption $\Delta E_{av}/k = B$ to be satisfactory.

The temperature coefficient of resistance $\partial R_f / \partial T \simeq R_f / \Delta T$(ohms/°C), as determined from the curve in Figure 1 at $H_0 = 0$, drops from 3.5 to 2.5% with an increase in ferrite temperature from 50 to 150°C. We will note that the temperature coefficients of resistance of certain ordinary semiconductor thermistors have the same values.

For investigation of the dependence of ferrite resistance on the constant magnetizing field a cylindrical specimen was magnetized by a longitudinal field of $H_0 = 3500$ oersteds. In this case the ferrite resistance at room temperature decreased by 5%, which is seen from

Figure 2 showing the temperature dependence of the longitudinal galvanomagnetic effect of a manganese ferrite with H_0 = 3500 oersteds. It is seen from Figure 2 that the longitudinal galvanomagnetic effect for manganese ferrite is negative in sign. This result agrees with the data given in reference [3]. Upon increasing the temperature of the ferrite specimen to 100°C and upward to the Curie point the longitudinal galvanomagnetic effect becomes somewhat small (a change of $R_f < 1\%$ at operating H_0 fields). This permits us to disregard the galvanomagnetic effect in calculations for operation at high power levels. Determination of $R_f = f(T)$ and measurement of the galvanomagnetic effect were performed by the methods described in reference [3] (with a few insignificant changes).

The static volt-ampere characteristics of the ferrite with $H_0 = 0$ and $H_0 = 3500$ oersteds are shown in Figure 3. It is quite evident from these characteristics that the ferrite is a nonlinear conductor. At small currents an increase in current results in an increase in voltage drop at the ferrite. As the current through the ferrite is further increased the voltage begins to decrease due to nonlinearity caused by heating of the ferrite. These nonlinear properties of the ferrite are inertial. The volt-ampere characteristic is required for selection of the operating portion on the curve of $U_f = f(I_f)$ and design of the bridge circuit.

Shape of the static volt-ampere characteristic is quite similar to that for ordinary semiconductor thermistors. The curve for $R_f = f(P_f)$ at room temperature is shown in Figure 4, from which it is seen that with an increase in power absorbed ny the ferrite at ferrimagnetic resonance there is a decrease in ferrite resistance.

The relation $R_f = f(P_f)$ is described by formula (3). With known A and B, coefficient c is easily determined from the curve in Figure 4. From this same curve we may determine sensitivity of the ferrite, which is characterized by the change in resistance R_f in ohms with a change in P_f of 1 mw and is defined by the formula

$$\eta_f = \frac{\partial R_f}{\partial P_f} \simeq \frac{\Delta R_f}{\Delta P_f} \text{ ohms/mw}$$

In the given ferrite specimen the sensitivity in the central region of the curve for $R_f = f(P_f)$ at room temperature is 1400 ohms/mw.

For the sake of comparison we will note that the sensitivity of an instrument thermistor of type T8M is 66-90 ohms/mw at the operating point.

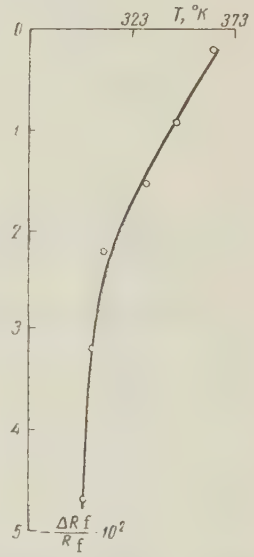


Figure 2. Temperature dependence of longitudinal galvanomagnetic effect of manganese ferrite with a magnetizing field H_0=3500 oersteds.

$R_f = f(P_f)$ was plotted from measurements performed with the setup described in detail in reference [4]. One difference lay in the fact that for measurement of R_f a UM-3 meter was connected to the ferrite through an opening in the narrow wall of a rectangular cavity. In determining P_f it was considered that all the energy of the cavity is absorbed by the ferrite at ferrimagnetic resonance.

Thus, the presented experimental and design data characterize the ferrite as a ferrimagnetic thermistor and qualitatively confirm the possibility of using ferrites for the absolute measurement of microwave power.

The principal difference between a ferrimagnetic thermistor and an ordinary thermistor is that the former interacts only with magnetic field of a microwave oscillatory system and the latter interacts with the electric field.

Electric breakdown between particles cannot occur within the ferrite specimen and with appropriate location of the specimen in the oscillatory system the dielectric strength of these particles is not decreased. This permits use of the ferrimagnetic thermistor for measurement of high-level microwave power. Another distinctive feature of the ferrimagnetic thermistor is the possibility of regulating the sensitivity at the operating point

by changing the value of the magnetizing field. This property of the ferrimagnetic thermistor permits widening the range of measurable powers with the given specimen.

By virtue of its properties the ferrimagnetic thermistor is, as it were, "self-protected" against overloads. This is due to the fact that with an increase in temperature of the ferrite specimen $M_0 \to 0$ and, consequently $P_f \to 0$.

This immunity to overloads is the most important advantage of the ferrimagnetic semiconductor thermistor.

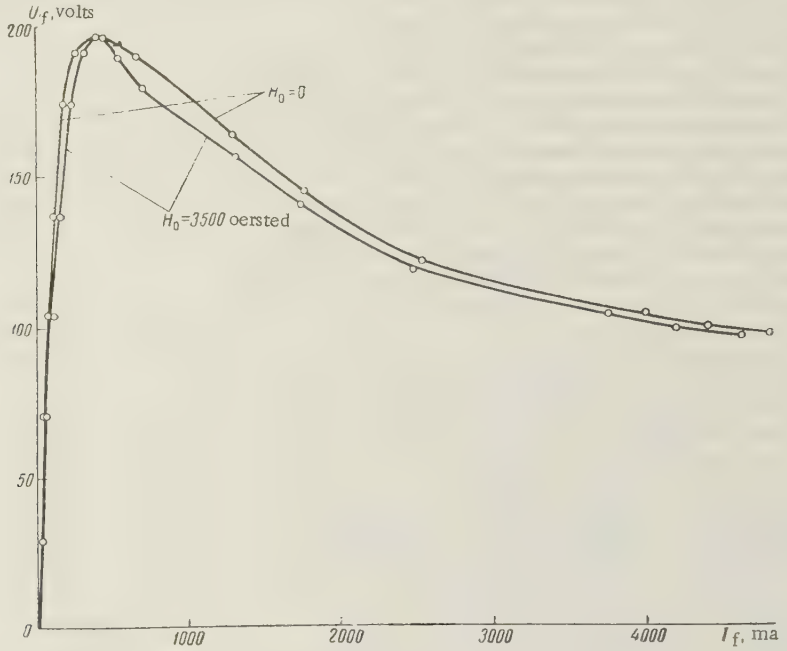


Figure 3. Static volt-ampere characteristic of manganese ferrite for different values of magnetizing field.

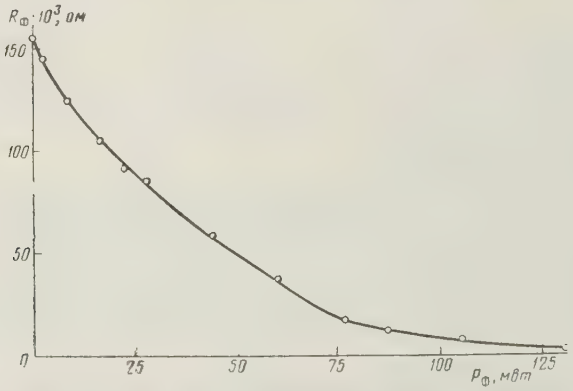


Figure 4. Dependence of resistance of manganese ferrite on power absorbed at ferrimagnetic resonance and room temperature.

I wish to thank Ya. A. Monosov and V.I. Pronenko for their critique of the results of this work.

REFERENCES

1. K. P. Belov, A.A. Popova, Ye.V. Talalayeva, Electric and galvanomagnetic properties

of manganese monocrystals, Kristallografiya, 1958, 3, 6, 734.
2. R.A. Valitov, V.N. Sretenskiy, Radioizmereniya na sverkhvysokikh chastotakh, [Radio measurements at microwavelengths], Voyenizdat, 1951.
3. K.P. Belov, Ye.V. Talalayeva, Temperature dependence of the galvanomagnetic effect and electrical resistance in manganese ferrite in the poly- and monocrystal state, NDVSh (Fiziko-matematicheskiye nauki), 1958, 2220.
4. Ya. A. Monosov, A.V. Vashkovskiy, Experimental investigation of nonlinear phenomena in ferrites at microwavelengths, Radiotekhnika i elektronika, 1960, 5, 1, 106.

Submitted to the editors 29 September 1960

INFLUENCE OF TEMPERATURE PREHISTORY ON PERMEABILITY OF NICKEL-ZINC FERRITES

D.D. Mishin and L.A. Kalyagina

INTRODUCTION

The combined influence of thermal and magnetic history on intensity of magnetization of metallic ferromagnets in weak and medium fields (magneto-thermal hysteresis) has been investigated in detail by Shur and colleagues (reference [1]) and by Kirenskiy and colleagues (reference [2]). It has been shown that magneto-thermal hysteresis depends substantially on the intensity of a magnetic field acting constantly during a temperature cycle, the range of the temperature cycle, the crystallographic direction, etc. Shur and Baranove (reference [3]) have shown that the magneto-thermal hysteresis may be explained by irreversible processes of displacement of the boundaries between domains under the influence of the magnitizing field with a change in temperature of the ferromagnetic material. L.V. Kirenskiy and colleagues (reference [2]) explain magneto-thermal hysteresis as irreversible processes of reorganization of magnetic structure occurring in a definite region of the magnetizing fields. It follows from the above cited works that magneto-thermal hysteresis is somewhat complex phenomenon.

In order to investigate a few of the important aspects of magneto-thermal hysteresis it is necessary to simplify discussion of the factors affecting the ferromagnetic material (for example, to exclude the effect of the magnetic field and to study the influence of the temperature prehistory on the susceptibility of the ferromagnetic material. In this special case (developed by us in reference [5]) the magneto-thermal hysteresis may be referred to as the thermal hysteresis of the magnetic structure, which can be observed from the influence of the temperature prehistory on the permeability of the ferro- or ferrimagnetic material. The present report describes the influence of temperature prehistory on the permeability of a nickel-zinc ferrite of class F-1000 in the presence of constant and alternating fields of different intensities.

1. SPECIMENS AND MEASUREMENT PROCEDURES

In order to decrease the influence of the earth's magnetic field the specimens of nickel-zinc ferrite (F-1000) were made in toroidal shape. Dimensions: outer diameter 22 mm; inner

diameter 10 mm; height 7 mm.

Measurement of the permeability was performed by the bridge method with variable capacitance at 4.5 kc. The output of a ZG-10 audio oscillator was applied to the bridge. The null indicator was a measuring amplifier of type 28IM with narrow passband. Heating and cooling of specimens with windings were achieved in an oil medium at an experimentally predetermined rate which

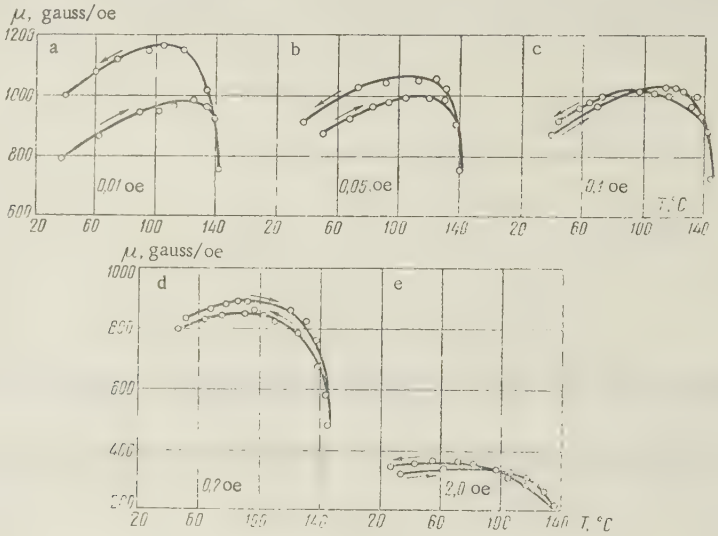
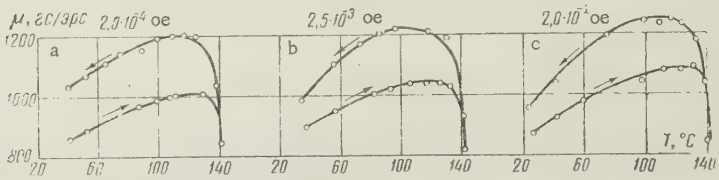


Рис. 1



ensured thorough heating of the specimens. Permeability measurement was performed with slow heating from 20°C to Curie temperature (150°C) and subsequent slow cooling. The number of turns of the measuring winding was 40 and of the magnetizing winding 25. The effective value of magnetic field intensity at which the permeability was measured varied from $2.0 \cdot 10^{-4}$ to $2.0 \cdot 10^{-2}$ oersteds. The magnetizing winding was fed from a high-resistance dry-cell battery. The magnetizing field varied from 0.01 to 2.0 oersteds

In order to establish the influence of the number of heating and cooling cycles on the thermal hysteresis of permeability the latter was measured in the same specimen during each of twelve such cycles.

2. EVALUATION OF MEASUREMENTS

Figure 1 shows the results of measurement of the permeability of F-1000 ferrite during heating and subsequent cooling with different magnetizing field intensities. The magnetizing field intensities are indicated in the figure. As is seen from Figure 1a, maximum hysteresis occurs with the weakest magnetizing field (0.01 oe). The hysteresis decreases as the magnetizing field increases. At weak magnetizing fields of 0.01 and 0.05 oe (Figure 1a and 1b) the curve for cooling passes above the curve for heating. With a magnetizing field of 0.1 oe (Figure lc) the curves for cooling and heating intersect. With a magnetizing field of 0.2 oe (Figure 1d) the curve for cooling passes below the curve for heating, whereas with a magnetizing field of 2,0 oe (Figure 1e) both curves (under the given conditions of measurement) practically merge. Thus it is seen from Figure 1 that in the given case a constant

magnetizing field only decreases the hysteresis. As follows from reference [4], magnetizing fields with intensities of up to 0.2 oe are fields in which over the entire temperature range only reversible magnetizing processes occur, whereas fields with intensities of the order of 2.0 oe are fields of reversible and irreversible processes. Thus, regardless of whether the magnetizing processes in the region of the given magnetizing fields are reversible or irreversible, the hysteresis only decreases as the field intensity increases.

Figure 2a and 2b show similar curves -- the measurement cycles of permeability with heating and subsequent cooling of an F-1000 specimen without a magnetizing field but with different effective values of magnetic field intensity (indicated in Figure 2) at which the bridge of the measuring circuit is balanced. It is seen from these curves that with a considerable change in effective value of the measurement field intensity (from $2 \cdot 10^{-4}$ to $2 \cdot 10^{-2}$ oe) the permeability hysteresis changes inconsiderably.

The table lists the results of measurement of the permeability hysteresis of the same specimen during the first through the fifth cycles and during the twelfth cycle of heating and cooling.

It is seen from the table that the hysteresis decreases with each cycle, but as the number of cycles increases this decrease becomes smaller.

Cycle No.	$\mu_1 - \mu_2$	Cycle No.	$\mu_1 - \mu_2$
1	160	4	90
2	110	5	90
3	100	12	80

Note: μ_1 is the permeability measured at 100° C during heating; μ_2 is the permeability measured at 100° C during cooling.

The observed hysteresis is apparently explained by the thermal hysteresis of the domain structure of the ferrits, which undergoes an irreversible change upon heating and cooling. Irreversible changes in magnetic structure of a ferrite upon heating and cooling may occur both during irreversible processes of displacement of domain boundaries (reference [3]) and during processes of reorganization of the domain structure of the ferrite (reference [2]).

REFERENCES

1. Ya. S. Shur, V.I. Drozhzhina, ZhETF, 1947, 17, 6, 607; Ya.S. Shur, N.A. Baranova, ZhETF, 1950, 20, 2, 183.
2. L.V. Kirenskiy, D.A. Laptey, A.I. Drokin, R.P. Smolin, Fizika metallov i metallovedeniye, 1960, 9, 3, 337.
3. Ya. S. Shur, N.A. Varanova, Dokl. ANSSSR, 1954, 94, 5, 825.
4. D.D. Mishin, L.D. Drobchenko, Izv. vuzov MVO SSSR (Fizika), 1960, 4, 131.
5. D.D° Mishin, N.G. Plantus, E.E. Admova, 3-ye Veseovyznoye soveshchaniye po fizike, fiziko-khimicheskim svoystvam ferritov i fizicheskim osnovam ikh primeneniya (Tezisy dokladov) [Third All Union Conference on the Physics and Physico-Chemical Properties of Ferrites the Physical Principles of their Application (Reports]; Minsk, 1959.

Submitted to the editors 22 August 1960

ENERGY DISTRIBUTION OF FIELD-EMISSION ELECTRONS IN SEMICONDUCTORS

I. M. Bronshteyn and Ya. M. Shchuchinskiy

1. It was shown in reference [1] that the shape of the curve for energy distribution of slow secondary electrons varies with the thickness of the spray coating of Ba on W. Instead of the single maximum which is characteristic of W, there appear two maxima. With an increase in the thickness θ of the Ba layer the relative value of the first ("tungsten") maximum decreases and that of the second ("barium") increases. With $\theta \simeq 10$ atomic layers of Ba on W the first maximum disappears and the curve for energy distribution of slow secondary electrons acquires the form characteristic for a thick layer of Ba. In this same work (reference [1]) it was also pointed out that the second (barium) maximum upon adsorption of Ba atoms onto W is displaced in the direction of lower energies, following the change in effective work function of the target, wherein its position is finally fixed at a $\theta \simeq 2$ atomic layers of barium.

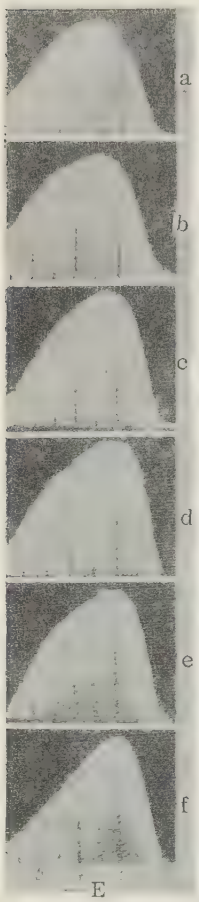


Figure 1. Change in low-voltage maximum of curve of energy distribution of secondary electrons according to thickness of Be coating on W.

a, W; b, $\theta \simeq 0.3$ atomic layers; c, 1 atomic layer; d, 2 atomic layers; e, 6 atomic layers; f, e-- 12 atomic layers.

This report presents the energy spectra of slow secondary electrons of thin layers of Be adsorbed onto W. The shape and position of the maximum of the distribution curves for slow secondary electrons of Be and W differ considerably although their work functions are almost the same ($\varphi_{Be} = 5$ev, $\varphi_W = 4.52$ ev).

The device, techniques and procedure of the experiment were described in detail in reference [1]. We will note that the curve $\sigma(E_p)$ for Be, obtained in vacuum at $5 \cdot 10^{-9}$ mm Hg, is in excellent agreement with the curve obtained in a vacuum of $\sim 10^{-7}$ mm Hg (reference [2]). Thus, the vacuum conditions ($p \simeq 10^{-8}$ mm Hg) under which our measurements were performed appear to be acceptable.

2. Figure 1 (a-f) shows the dynamics of the change in curves of energy distribution of slow secondary electrons according to the thickness of the spray coating of Be on W. It is seen from these curves that with an increase in θ the position of the maximum and the shape of the curve vary. The position of the maximum, as in the adsorption of Ba, is fixed at $\theta \simeq 2$ atomic layers; however, the relative contribution of electrons with energies of >3-3.5 ev is still quite large. With a further increase in θ the half-width of the curve decreases and with $\theta > 12$ atomic layers (Figure 1f) the curve of energy distribution of slow secondaries does not differ from that for a thick layer of Be.

Thus the sublayer ceases to have any effect on the shape of the curve of energy distribution of slow secondaries at a Be thickness of $\theta \simeq 12$ atomic layers that is with $\theta = \lambda$ {(λ is the upper limit of the free path of slow secondaries in Be (reference [3])}.

While the change in work function, as shown in reference [1] affects the energy spectrum

of secondary electrons, it evidently follows from the above data that the position and shape of the maximum in the given case are determined not only by the work function but also by other emitter properties.

REFERENCES

1. I.M. Bronshteyn, Ya. M. Shchuchinskiy. Radiotekhnika i elektronika 1960 5 10. 1650.
2. M.G. Nakhodkin, V.O. Romanovskiy, Ukr. fiz, zh., 1954, 4, 3, 479.
3. I.M. Bronshteyn, R.V. Segal', Fizika tverdogo tela, 1959, 1, 10, 1489.

Submitted to the editors 11 June 1960.

ENERGY SPECTRUM OF SLOW SECONDARY ELECTRONS IN ADSORPTION OF BERYLLIUM ONTO TUNGSTEN

A.G. Zhdan and M.I. Yelinson

Previous reports (references [1, 2]) have presented theoretical and experimental investigations of the influence of an internal electric field on the field emission of semiconductors

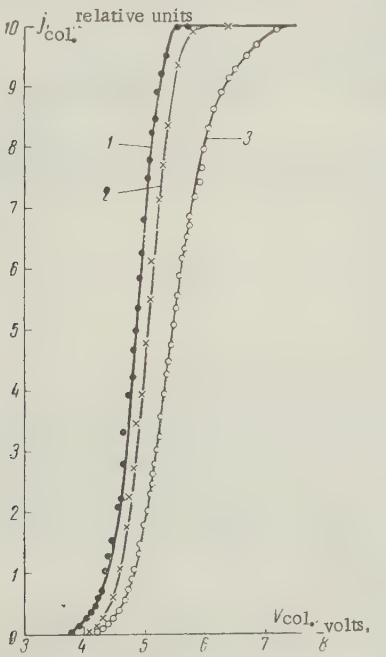


Figure 1. Delay as a function of current density j at room temperature:
1, j = 66a/cm^2; 2, j = 125 a/cm^2; 3, j = = 362 a /cm^2.

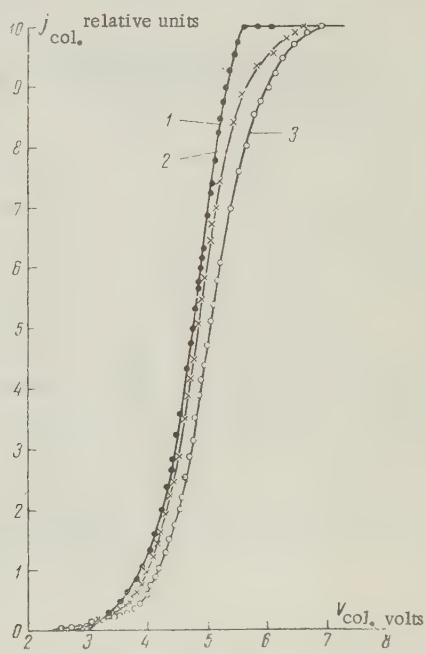


Figure 2. Delay as a function of current density j at 1100° C;
1, j = 15 a/cm^2; 2, j = 187 a/cm^2; 3. j = = 494 a/cm^2.

and have shown that in the presence of sufficiently strong fields the concentration of electrons in the conduction band and their energy distribution change significangly in comparison with the case of absence of an internal field, as a consequence of which an essentially nonequilibrium emission is attained.

The character of the change in the energy distribution function of field-emission electrons is best seen by measuring the spectrum of field-emission electrons at different selected currents.

We measured the spectra for thin semiconductive layers of quartz on carbon-activated tungsten (for the method of obtaining the layers and for their properties see reference [2]).

The experimental device was similar to that described in reference [3] with the exception of the method of adjustment and the collector material. The strong focusing of the electron stream and the elimination of stray lens effects resulted in excellent resolution.

Figure 1 shows the delay curves at room temperature for three values of mean density of current from a field-emissive point. From the figure it is seen that the delay curve expand from ~ 1.75 to ~ 3.25 ev with a change in j from 66 to 362 amp/cm.

With an increase in current density in several cases there was observed an expansion of the curves up to ~ 7 ev. Figure 2 shows the delay curves for three values of j at 1100°C. The nature of the change in the curves is somewhat different: expansion is noticeable even at small values of j, but there is little change in the width of the curve. The position of the delay curves on the V_{col} axis permits estimation of the voltage drop at the SiO_2 +C layer and, consequently, of the mean value of the internal field E_{int} in the layer. Considerable expansion of the spectrum was observed with $E_{int} \simeq 10^4$ v/cm.

The obtained results permit the following important conclusions.

1. The essentially nonequilibrium nature of field emission of high-resistance semiconductors is confirmed.
2. With internal fields E_{int} 10^4 v/cm there occurs a considerable increase in electron temperature.
3. With an increase in lattice temperature the increase in electron temperature is retarded.
4. In the presence of strong internal fields a considerable number of emitted electrons pass over the potential barrier.

REFERENCES

1. M. I. Yelinson, Radiotekhnika i elektronika, 1959, 4, 1, 140.
2. M. I. yelinson, G. F. Vasil'yev, A. G. Ahdan, Radiotekhnika i elektronika, 1959, 4. 10, 1718.

Submitted to the editors 3 January 1961

LETTERS TO THE EDITORS

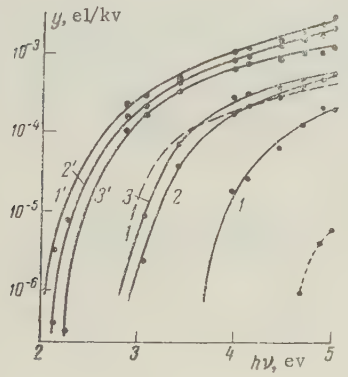
PHOTOELECTRON EMISSION OF LaB_6

L.S. Miroshnichenko

According to data in the literature, an interesting property of LaB_6 is its metallic conductivity and low work function (reference [1]) in comparison with ordinary metals. Since in the latter the work function is decreased by means of the adsorbed BaO molecules to a value of the order of 2-3 ev, it is of interest to investigate the possibility of such a decrease and a corresponding shift of the red threshold of photoemission in lanthanum hexaboride. The high sensitivity of LaB_6 to surface poisoning by residual gases is also known (references [3,4]). According to our observations the difference in work functions of a cleaned surface and one located in an atmosphere of residual gases at $p \simeq 1 \cdot 10^{-8}$ mm Hg is 0.8 ev. Our experiments were conducted with specimens with different surface conditions.[1]

The experiments were performed in sealed tubes of a design previously described (reference [2]), which allowed heating of the specimens by electron bombardment. During careful evacuation of the device the specimens were heated at 1200°C for approximately 20 hours.

The results are shown in the figure, wherein the spectral characteristics of quantum yield of the emitted particles correspond to the different states of the LaB_6 surface.



Curve 1 is for the initial surface state of a specimen kept for a long time in an atmosphere of residual gases at $p \simeq 1 \cdot 10^{-8}$ mm Hg. Surfaces purified to a different degree (by heating to 1100° C) are represented by curves 2 and 3. The long dashed curve represents Lafferty's data (reference [3]). The corresponding work function values obtained from our data by Fowler's method are 3.55, 2.89, and 2.77 ev. The optimum coatings of these initial surfaces with molecules of barium oxide lead to a shift of the spectral characteristics to positions 1', 2', and 3' and to different values of work function (2.07, 2.12, and 2.24 ev, respectively). The best purified surface gave the lowest optimum decrease in work function. This decrease (0.53 ev) is extremely small in comparison with the decrease in work function of such ordinary metals as tungsten, gold, tantalum, et al..

Experiments with GdB_6 specimens in our arrangement led to results contrary to those given in the literature. The spectral charateristic for the purified surface of our specimen is given in the figure as a short dashed curve, from which the work function is found to be approximately 4.6 ev instead of the value of 2.1 ev given in the literature (reference [1]). With the use of Ba) the work function of our specimen was reduced by slightly more than 2 ev.

I wish to thank P. G. Borzyak for his constant interest and assistance in this work and also G.V. Samsonov and Yu. V. Paderno for submitting the specimens for the investigation.

REFERENCES

1. G.V. Samsonov, Uspekhi khimii, 1959, 28, 189.
2. P. G. Borzyak, O.G. Sarbey, ZhTF, 1958, 28, 1905.

3. J. M. Lafferty, J. Appl. Phys., 1951, 22, 299.
4. R. W. Decker, D. W. Stebbnis, J. Appl Phys., 1955, 26, 1004.

Physics Institute AN USSR

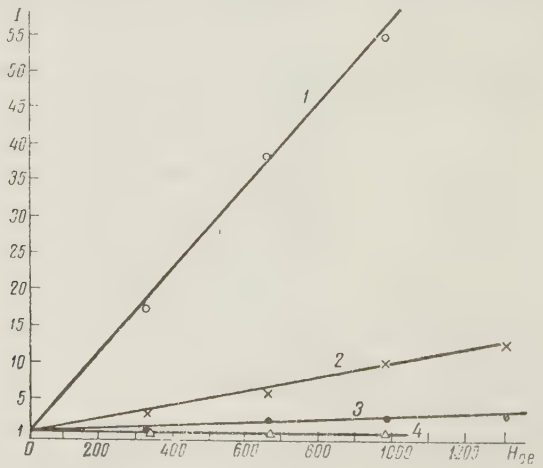
Submitted to the editors 9 November 1960

RECOMBINATION RADIATION OF CESIUM PLASMA IN A MAGNETIC FIELD

Yu. M. Aleskovskiy, V. L. Granovskiy and Ye. Mikhalets

Upon the application of a longitudinal magnetic field to the positive column of a low-pressure discharge there is a decrease in the diffusion of electrons and ions toward the walls in a direction perpendicular to the field. Hence the mean life of charge carriers in the plasma in increased. The ion balance is mean life of charge carriers in the plasma is increased. The ion balance is maintained at a lower ionization frequency (z_1) at one electron per second, hence intensity of the longitudinal electric field (E_z) and the electron temperature (T_e) in the stationary plasma are reduced (reference [1]), which was observed experimentally in reference [2].

In this connection we may assume that within the magnetic field the contribution of charged particles leaving the plasma due to volume recombination is increased (reference [3]). The absolute number of recombinations per cm^3 per second must also increase somewhat. [1] However, spectroscopic observations of the recombination radiation of a discharge (Davis, reference [5]) have not agreed with the above conclusion.



Intensity of recombination radiation as a function of magnetic field intensity: 1, p=8.2μ Hg; 2, p=18 μ Hg; 3, p=36μ Hg; 4, p= =74 μ Hg

[1] As is shown in reference [4], in a decaying plasma the absolute number of recombinations m rise sharply with an increase in magnetic field.

We undertook the study of the influence of a magnetic field on electron recombination radiation, in a low-pressure stationary discharge in cesium vapor. For this purpose we measured the intensity of the recombination continuum with a limit of 4943° A, corresponding to electron capture at the $Cs6P_{1/2}$ level (reference [6]).

A discharge tube with diameter of 25 mm was placed within a uniform magnetic field created by two solenoids. Through the gap between the solenoids and at the center of the tube the discharge radiation was led out to a monochromator and then recorded by means of a photomultiplier or observed visually. The discharge current varied from 1 to 2.4 amp and the cesium vapor pressure varied from 2 to 130 microns Hg.

It was established that the magnetic field has a strong influence on radiation of the positive column. Intensity of all emission lines of cesium decreases with an increase in field. For example, the intensity of the Cs 4555-4593 A doublet decreases by 5-10 times with $H \simeq 1000$ oersteds. The recombination glow, on the other hand, increases with the magnetic field. This effect is particularly evident at low pressures. At the lowest pressure the recombination glow is not observed in examining even a well adapted gas through the spectroscope, but application of a field of the order of 100 oersteds leads to the appearance of a clearly visible continuum.

The figure shows the intensity of recombination radiation as a function of the magnetic field at different pressures of Cs vapor.

In all curves it was assumed that $I = 1$ when $H = 0$. In the calculations a correction was introduced for that portion of the light of a bright Cs $8^2P_{1/2,3/2} - 6^2S_{1/2}$ doublet which was scattered in the monochromator optical system. As was to be expected, with an increase in pressure the influence of the magnetic field on the recombination glow of the plasma decreases, for there is a decrease in the mean free path of electrons and, consequently, in the quantity $\omega\tau$ determining the effect of the magnetic field on the electron parameters of the plasma.

REFERENCES

1. L. Tonks, Phys. Rev., 1939, 56, 360.
2. R.J. Bickerton, A. Engel, Proc. Phys. Soc. B, 1956. 69, 468.
3. I.A. Vasil'yeva, Radiotekhnika i elektronika, 1960. 5. 12. 2015.
4. A.S. Sirgiy, V.L. Granovskiy, Radiotekhnika i elektronika, 1960, 5, 9, 1522.
5. L.W. Davies, Proc. Phys. Soc. B, 1953, 66, 33.
6. F.L. Mohler, Phys. Rev, 1928, 31, 187.

Physics Faculty
Moscow State University
imeni M.V. Lomonosov

Submitted to the editors 5 January 1961

CHRONICLE

SYMPOSIUM ON WAVE DIFFRACTION

From 26 September to 1 October there was held in Odessa a symposium on diffraction of waves. The symposium was organized by the Committe on Acoustics AN SSSR conjointly with the Acoustics Institute AN SSR and the Odessa Electrical Engineering Institute of Communications. The sessions were held on the premises of the latter institute. The theme of the symposium was extremely broad and embraced waves of the most varied nature, which permitted extensive interchange of views between the scientific workers in various specialties during the discussion of common problems.

The plenary sessions were devoted chiefly to summary reports. At the first plenary session after the initial opening of the symposium and the introductory address by V.A. Fok, G.A. Grinberg gave a review of the work performed at the mathematical physics division (which he heads) at the Leningrad Institute of Physics and Technology AN SSSR on the theory of diffraction of electromagnetic waves and the development of general methods of mathematical physics.

At subsequent plenary sessions reports were given by V.A. Fok — "Certain electrodynamic problems for a hollow cylinder of finite length" (concerning methods of numerical solution of boundary problems in electrostatics and electrodynamics for a hollow conductor) — and Ye. L. Feynberg — "Diffraction problems in the physics of elementary particles" (concerning diffraction problems arising in the scattering of nucleons at the nucleus and mutual scattering of protons, electrons and π-mesons, and also in inelastic scattering of particles). L.A. Vaynshteyn's report on "Electromagnetic deffraction and boundary problems" gave a summary of works performed in the USSR during the years 1957-1960 and expressed the opinion that in the near future diffraction theory will expand rapidly, on the one hand, due to the devlopment of numerical methods on the basis of modern computer techniques and, on the other hand, due to asymptotic methods which provide approximate and graphic solutions of diffraction problems in "quasi-optic" areas where direct numerical methods are ineffective. G.D. Malyuzhints' report "The ideas of Thomas Young and asymptotic laws of diffraction" presented an asymptotic description of diffraction by means of a parabolic equation given in radial coordinates and considering the transverse diffusion of wave amplitude; the report also discussed the prospects for development of analytical and machine calculating methods of solving diffractional problems.

In the summary report by A.S. Alekseyev, V.M. Babich and B. Ya. Gel'chinskiy, "Radial method of calculating the intensity of wave fields" the mathematical problems associated with this method were discussed. N.N. Monseyev's report, "The present state and problems of the nonlinear theory of surface waves, " dealt with hydrodynamic waves.

The report presented by L.D. Bakhrakh and A.A. Pistol'kors on "Certain problems of diffraction theory in antenna techniques at centimeter wavelengths" stated a number of diffraction problems (diffraction of compound waves at a small mirror, radiation of antennas located at a conducting cone, diffraction and radiation of surface waves, influence of antenna blister of conical shape on antenna operation, etc. M.D. Khaskind's report on "Certain problems in the diffraction and excitation of waves at an impedance plane" showed that for bidimensional problems in the diffraction and radiation of hydrodynamic surface waves (these problems are associated with the hydrodynamics of seagoing vessels) the mathematical statement is the same as for acoustic and electrodynamic problems in the diffraction and excitation of waves above an impedance plane, and presented rigorous and approximate solutions of a number of problems of this type.

Sectional sessions were held simultaneously in four sections.

A. Rigorous and numerical solutions of boundary diffraction problems. Asymptotic metho in boundary diffraction problems.

B. Nonstationary problems. Rayleigh waves. Waves in heavy liquid.

C. Waves in lamellar media. Grids and corrugated surfaces. Propagation of waves.

D. Regular and periodic waveguides. Irregular waveguides. This section held an additional session devoted to diffraction in optical devices.

Since it is not possible to list all the section reports and communications, we shall mention only a fiew of them with the interests of the readers of our journal in mind.

In the report by G. D. Malyzhints and A. A. Tushilin on "Electromagnetic field excited by an electric dipole in a wedge-shaped region with ideally conducting boundaries" was given a rigorous solution of the problem in the form of Sommerfield integral and simple asymptotic formulas were derived for a field at great distances. In his work in "Integral equations of antenna theory" N. N. Govorun investigated integral equations for current density at an antenna represented as a rotational body with impedance surface and presented the numerical results for a thin cylindrical dipole obtained by high-speed computer. The theory of dipoles was also discussed in P. Ya. Ufintsev"s report on "Scattering of plane waves at thin cylindrical conductors," wherein the results for current waves in thin cylindrical conductors were obtained by the method of slowly varying functions used (in conjunction with simply physical consideration) for calculation of the scatter characteristics of passive dipoles.

It is known that by solving the problem of diffraction of unit step or a nonstationary wave of any mode it is possible not only to construct a formal solution of the problem of diffraction of a monochromatic wave, but also to determine (effectively) the short-wave asymptote of the latter problem. In V. A. Borokov's report on "The tridimensional problem of diffraction of a plane wave at a plane shield with wedge-shaped notch" the nonstationary problem of diffraction of a plane scalar wave at a plane with notched angle $\alpha(\pi/2 < \alpha < 3\pi/2)$ is reduced to the Dirichlet problem for the Laplacian in a bounded tridimensional region. Numerical solution of this problem permits determining the asymptotic properties of a monochromatic diffraction field. In the report by A. Ya. Povzner and I. V. Sukharevskiy "On asymptotic expansions in certain problems in the theory of diffraction of short waves" there is investigated an infinite region with an ideally reflecting and sufficiently smooth boundary wholly "illuminated" from the source. The asymptotic formulas correspond to the laws of geometric optics.

Similar results for a more special case and another method are obtained in the communication by B. Ye Kinber on "Diffraction of a cylindrical wave at the inside of a circular cylinder," where the asymptotic formulas contain multiply reflected rays, caustics and "whispering gallery" waves traveling along the inside of the cylinder. In a second communication on "Approximate solution of the problem of diffraction at a parabolic mirror of finite dimensions" B. Ye. Kinber showed that in a logical theory of mirror antennas, based on the concept of diffraction rays, it is necessary to consider the specific aspects of a concave surface, particularly the multiple reflections of the fringe wave.

In the reports "Scattering of plane and cylindrical waves at an elliptical cylinder and the concept of diffraction rays" (L. A. Vaynshteyn and A. A. Fedorov) and "Short-wave asymptote to Green's function in the problem for parabolic surfaces" (V. I. Ivanov) the asymptotic laws of diffraction at convex, ideally reflecting bodies (elliptical cylinder, paraboloid of rotation) are investigated.

In the work by Ye. N. Mayzel's and P. Ya Ufimtsev on "Reflection of electromagnetic waves of circular polarization from metallic bodies" it is shown that there may be distinguished in a scatter field (both theoretically and experimentally its "nonuniform" component, caused by deformation (break) of the surface. The report by L. M. Brekhovskikh and V. A. Yeliseyevnin "On propagation of waves in a nonuniform waveguide" is devoted to the propagation of sound waves in a lamellar medium of special type forming a waveguide, the properties of which vary along the path of propagation. In "The propagation and reflection of radio waves from an ionized layer" (G. I. Makarov) and "New asymptotic representations of the Whittaker functions for problems of propagation of radio waves" (E. M. Gyuninen and G. I. Makarov) there is given asymptotic treatment of propagation of radio waves in a lamellar atmosphere with $\vec{E}$ grad $\epsilon \neq 0$. Tropospheric propagation of radio waves is discussed in the report by L. M. Ponomarenko, "The role of coherent scattering in long-distance propagation of ultrashort waves."

Investigation of the passage of sound waves or electromagnetic waves through an infinite periodic grating formed by parallel cylinders is a complex diffraction problem even with the condition that the period is small in comparison with the wavelength. Although in the latter case electromagnetic waves) the wave field in the vicinity of the grating plane must change to static fields, the corresponding "interlacing" was until recently associated with definite difficulties. The theory of such gratings was discussed in the reports by M. I.

Gurevich ("Acoustic conductivity of a close-spaced grating), G. D. Malyuzhints ("Mean boundary conditions at planes defining a remote field in the diffraction of long waves at a dense, acoustically rigid grating) and A.N. Sivov ("Oblique incidence of a plane wave at a plane grating of parallel conductors"). In the report by A.M. Model' and N.V. Talyzin ("Diffraction of a plane wave at an infinite grating consisting of individual dipoles"), there was investigated the reflection of an electromagnetic wave from a system of dipoles located in a plane in alternating order.

The reports by M. D. Khaskind ("Propagation of electromagnetic waves above a gyrotropic medium" and "Radiation of electromagnetic waves above a thin gyrotropic layer) presented the solution of complex problems of excitation of waves by given sources in the vicinity of a magnetized ferrite or plasma.

Various problems in the propagation of radio waves over the earth's surface were discussed in the reports by Yu. K. Kalinin and A.D. Petrovskiy (" Approximate boundary conditions and the diffraction of radio waves"), Yu.K. Kalinin ("Diffraction of radio waves at surface discontinuities (a comparison of theory and experiment)"), V.V. Novikov, ("Propagation of pulse signals over a plane, uniform earth surface (nonstationary problem of Sommerfield)"), E.M. Gyuninen, G.I. Makarov and A.V. Manankova ("Electromagnetic-pulse propagation over the surface of a spherical earth"), G.N. Krylov ("Field of a vertical dipole and antenna over an earth with infinite conductivity"). In his report on "Propagation of electromagnetic pulses in a conducting medium" Ye. B. Khanakhbey investigated propagation in an infinite conductor and in a conducting medium bounded by a dielectric.

The theory of propagation at corrugated surfaces was discussed in reports by I. A. Urusovskiy ("Diffraction of sound at periodic irregular and nonuniform surfaces"), R.G. Barantsev ("Scattering of a plane wave at an arbitrary periodic surface"), B.F. Kur'yanov ("Scattering of sound at a rough surface formed by two types of irregularity") and V.I. Aksenov ("Scattering of electromagnetic waves at periodic irregular surfaces with finite conductivity"). In Yu.P. Lysanov's communication on "Diffraction of a plane wave at a nonuniform surface with continuously varying properties" it was shown that if the derivative of the reflection coefficient reaches a maximum at a certain value of coordinate x at the reflecting plane, then from the corresponding line a cylindrical wave diverges (as from a notch or break).

Among the reports devoted to waveguides it is necessary to mention that of M.G. Kreyn and G. Ya. Lyubarskiy, "Toward a theory of passband of periodic waveguides, " wherein general evaluations are presented for the passband of waves subject to the scalar wave equation. N.A. Kuzimin presented a report on "Potentials and the variation principle of equations of electrodynamics described in the nonorthogonal curvilinear system of coordinates." In the report on "Nonstationary processes in the propagation of pulse signals in a circular waveguide" A.A. Kovtun and G.A. Makarov presented a method of calculation and numerical results dealing with the excitation and propagation of radio pulses in a waveguide (with ideal walls and with consideration of finite conductivity of walls.)

In his report on "The theory of irregular acoustic waveguides" B.Z. Katsenelenbaum spoke of extension of the method of transverse sections (previously developed for radio waveguides) to the case of acoustic waveguides. For waveguides with slowly varying parameters this method permits explicit solution, that is, it permits determination of the amplitudes of all waves scattered by a regular section. The report by A.G. Sveshnikov and I.P. Kotik on "Methods of calculation of irregular waveguides" was devoted to methods whereby, by the use of high-speed computers, it is possible to perform calculation of an arbitrary, irregular waveguide. The report on "Synthesis of transmission lines from a given law of wave transformation" by Livshits and M. Sh. Flekser examined the general mathematical properties of the transmission matrix of a nonuniform section of line.

In conclusion it should be mentioned that the symposium had no precedent either at home or abroad and demonstrated the fruitfulness of a meeting of physicists, mathematicians and engineers dealing with the diffraction of various types of waves. At the final plenary session the symposium adopted the recommendation that another such conference be held in the spring of 1962 in Gor'kiy.

It is necessary to remark on the excellent organization of the symposium (representing much work on the part of the chairman of the organization committee, professor M.D. Khaskind of the Odessa Electrical Engineering Institute of Communications) and the hospitality with which the city of Odessa greeted the numerous participants coming to the symposium fro all over the Soviet Union.

VENEDIKT IVANOVICH KALININ

Soviet radiophysics has suffered a great loss. On October 15 1960, Venedikt Ivanovich Kalinin, Doctor of Physical and Mathematical Sciences, well-known in our country as a radiophysicist and honored scientist of the RSFSR, suddenly passed away.

The name of V.I. Kalinin is well known through his books and scientific research and for his activity as a teacher at Saratov University.

Venedikt Ivanovich Kalinin was born on March 16 1907 in Saratov into the family of a railroad employee. After finishing middle school he entered Saratov University, within the halls of which he passed almost his entire creative life from physics student to head of the chair of radiophysics created by him. While still a student under the direction of K. A. Leont'yev he began his research in the field generation of microwaves and in 1929 he published his first article, devoted to the regions of oscillations of the retarding-field generator.

Upon concluding his studies at the university, V.I. Kalinin worked for a while at the Central Radio Laboratory in Leningrad. Here he organized a group for study of decimeter waves. During this period he performed a number of important investigations in the generation of microwaves, in particular he developed a tube with a periodic grid, which was widely used at that time.

In 1933 V.I. Kalinin returned to Saratov and resumed work at Saratov University. Here his creative activity was particularly fruitful. This activity was extremely varied, attesting to his remarkable abilities and love of work. His wide scope and broad erudition permitted Venedikt Ivanovich to create the first series of valuable monographs on microwaves in our country. These included "Decimeter waves" (1935), "Decimeter and centimeter waves (1939) and "Generation of decimeter and centimeter waves"(1948). In addition to these works he conducted a number of original investigations. He published more than 50 scientific works, devoted chiefly to problems of the kinematic theory of microwave oscillators. A generalization of these investigations was made in his doctor's dissertation, defended in 1944.

The scientific interests of Venedikt Ivanovich Kalinin were closely linked with practical problems, Under his direction a number of assignments of great importance to the national economy were executed. Particularly trying work in this direction was performed during the war years, for which he was awarded a medal "For heroic service in the Great Patriotic War, 1941-1945."

Professor Kalinin devoted exceptionally serious attention to the teaching profession. He developed a number of exemplary radiophysics courses.

The lucid and absorbing lectures of Venedikt Ivanovich were used with great success in the auditorium. His extensive pedagogical experience was evident in a number of textbooks on which he, along with his students, expended much labor and effort. ("Introduction to Radiophysics," 1957; "A Radiophysics Practicum," 1956-1960).

Devoting great attention to creative scientific discussion, Venedikt Ivanovich organized in 1944 a permanent radiophysics seminar at Saratov University. This seminar has been instrumental in attracting young people to research work and conducive to a creative atmosphere in the department. Thereby Venedikt Ivanovich contributed to the training of scientific cadres of high skill, giving this work particular attention. Under his guidance more than ten persons have completed and successfully defended candidate dissertations.

Professor Kalinin was active in fields other than science and pedagogy. Kalinin took part in the actions of society, often presenting interesting lectures to the populace. One of his lectures was published in a separate brochure ("Russia — the birthplace of radio." 1949).

The fruitful activity of Venedikt Ivanovich Kalinin was highly esteemed: in 1959 he was awarded the title of honored scientist of the Russian Soviet Federated Socialist Republic.

Those around him knew Venedikt Ivanovich to be a man of high culture, straightforward and principled. The scope of his interests was extremely broad. He had a lively interest in literature and art, was attracted to creative photography, passionately loved Russian nature and was a zealous patriot. The shining example of Venedikt Ivanovich Kalinin will always be an inspiration to honorable service to Soviet science and our Motherland.

V. L. Ptarushev, G. M Gershteyn, V. Ya Krasil'nikov

RADIO ENGINEERING AND ELECTRONIC PHYSICS

The English Edition of Radio Engineering and Electronic Physics is mailed to subscribers within 18 weeks after the publication of the original Russian issue.

Russian electronic journals published by the
American Institute of Electrical Engineers
Translated by Royer and Roger, Inc.

	Subscription rates			
	Individuals		Libraries	
	$	£	$	£
Radio Engineering and Electronic Physics	28.50	10	57.00	20
Radio Engineering	14.25	5	28.50	10
Telecommunications	14.25	5	28.50	10

Royer and Roger translates and produces
the following Russian scientific journals:

Biophysics
Entomological Review
Geochemistry
Geodesy and Cartography
Izvestiya, Academy of Sciences of the USSR, Geologic Series
Pavlov Journal of Higher Nervous Activity
Problems of Oncology
Radio Engineering
Radio Engineering and Electronic Physics
Refractories
Sechenov Physiological Journal of the USSR
Soil Science
Telecommunications

Comments and inquiries regarding *Radio Engineering and Electronic Physics* and other translation journals should be sent to:

International Division
Royer and Roger, Inc.

1000 Vermont Avenue, N.W.
Washington 5, D.C.

41 East 28th Street
New York 16, New York

RADIO ENGINEERING AND ELECTRONIC PHYSICS

РАДИОТЕХНИКА И ЭЛЕКТРОНИКА

NUMBER 4 1961

CONTENTS